Volume 62, Part 1 February 1971

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by W. SLUCKIN

with J. ANNETT, W. S. ANTHONY, J. BROWN, R. CONRAD, N. F. DIXON, E. D. FRASER, J. A. GRAY, I. MARTIN, R. L. REID, H. R. SCHAFFER *and* P. B. WARR

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB

AMERICAN BRANCH:

32 EAST 57TH STREET, NEW YORK, N.Y. 10022

£2.50 net $8.50

Preparation of Papers for The British Journal of Psychology

In submitting papers for publication in *The British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this.
It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. 56 (1965) onwards.
A brief summary, suitable for use as an abstract, should precede the article.

Manuscripts should be submitted in duplicate.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation: and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.
Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.
The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.
Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–248.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.

RÉVÉSZ, G. (1951). Le probléme du gènie. *Année psychol.* **50**, 83–96.

World list abbreviations are used for journals.
The Oxford Dictionary is taken as the authority on spelling.

Subscriptions

The British Journal of Psychology is published quarterly.

Annual subscription (including postage) £7.50 net ($25.00 in the U.S.A.).

Second class postage paid at New York, N.Y.

Published by Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB and 32 East 57th Street, New York, N.Y. 10022.

Br. J. Psychol. (1971), 62, 1, pp. 1–11
Printed in Great Britain

32

DISCRIMINATION LEARNING AND THE STUDY OF TRANSFER IN YOUNG CHILDREN

By P. E. BRYANT*

Medical Research Council, Institute of Psychiatry, University of London

In discrimination learning experiments the location and dimension of distinctive cues are typically confounded. This raises the possibility that what has been described as dimensional transfer may be transfer of a response to a specific location within the visual display. This possibility was investigated in two experiments in which young children were subjects. In the first experiment it was shown that young children are able to transfer on the basis of the location of the relevant cues within the visual display. In the second experiment the strengths of dimensional and locational transfer were directly compared. Five-year-old children showed stronger locational than dimensional transfer, while 7-year-olds showed only dimensional transfer.

Recently considerable interest has been shown in the study of discrimination learning by young children, and the main focus of this interest seems to have centred on the question of dimensional transfer (Kendler & Kendler, 1962; Tighe & Tighe, 1966, 1968; Shepp & Turrisi, 1967; Wolff, 1967). Experiments on dimensional transfer appear to offer a simple and well-controlled way of investigating whether children's learning in one task is specific to that task or whether it is flexible enough to affect performance in a wide range of other tasks.

For example, if a subject learns a task in which red is correct and blue is incorrect, it may be that his learning is simply confined to these two colours, in which case it can be said to be specific. On the other hand, it may also be that he has learned to respond to or attend to any colour differences rather than to differences along other stimulus dimensions, such as shape or size. In the latter case learning could be said to be flexible, in that experience with one set of colours would affect behaviour in transfer tasks in which any colour differences were present.

Broadly speaking, two types of explanation have been offered for dimensional transfer when it has been found in children. One (Kendler & Kendler, 1962) attributes it to internal verbal mediation, the other (Zeaman & House, 1963) to a non-verbal attentional mechanism. Both hypotheses assume that, when dimensional transfer occurs in young children, it is based on a response to a stimulus dimension as such, and that, in fact, children are responding on the basis of an internal category such as colour or shape.

There is, however, a third possible explanation which is considered in this paper. This is that what is reported as dimensional transfer may actually be based on a response to the location rather than to the dimension of the relevant cues. Typically in a discrimination task the location and the dimension of relevant cues are confounded. For example, if stimuli differ both in colour and in shape, the colour cues come from the centre of the card, and the shape cues from its contour. If, after an initial colour task, a subject continues in a transfer task to respond to colour rather than to shape differences, this may be because he has learned that the central location of the stimuli provides the relevant cues, rather than that the relevant dimension is colour.

* Present address: Department of Experimental Psychology, University of Oxford.

There is, in fact, good evidence that rats (Ehrenfreund, 1948) and human adults (Kendler *et al.*, 1961) can record the location of relevant cues and transfer this information to a transfer task. However, the relevance of this response to dimensional transfer has not been explored. Nor is there evidence that young children can store and transfer information about the location of cues.

For these reasons two experiments were designed to investigate the response of young children to the location of cues. The aim of the first experiment was to investigate whether children are able to transfer information about the location of the relevant cues from one task to another. The aim of the second was to compare directly the relative strengths of locational and dimensional transfer in these subjects.

Experiment I

The aim of this experiment was to discover whether young children are able to transfer information about the location of relevant cues. The method adopted was to give the subjects two discrimination tasks in which different cues were used, but in both of which the cues came from one of two clearly defined locations within the stimulus. These two locations were the top and bottom halves of each card. For half of the subjects the location of the relevant cues was the same in both tasks. For the other half this location was changed between tasks. In the former case the relevant cues would be either in the top half in both tasks or in the bottom half in both tasks. In the latter case the relevant cues would be in the bottom half in one task and in the top half in the other. The crucial measure in this task was the speed of learning in the transfer task. If subjects for whom the relevant location is the same throughout learn this task more rapidly than those for whom it is changed, then this difference must mean that subjects have transferred a response to location from one task to the other.

The transfer task invariably contained cues in both locations, the cues in one location being relevant to the solution of the discrimination task, the cues in the other being irrelevant. However, in the initial task half the subjects learned with cues in the irrelevant as well as the relevant location. For the other half the cards only contained cues in one location, and not in the other, which was left blank. An example of such a task is one in which there are cues only in the top half of the cards, which is the relevant location, and no cues in the bottom half. The reason for the inclusion of this condition was to investigate the role of the irrelevant location on locational transfer. Work with rats (Turrisi *et al.*, 1969), with retardates (House & Zeaman, 1962), and with 5-year-old children (Bryant, 1967*a*) has shown that dimensional transfer is severely diminished by the absence of a variable irrelevant dimension in the initial task. If dimensional transfer is to be explained as a locational response, it follows that locational transfer should also be diminished after a task in which cues only come from the relevant location.

Thus this experiment had a 2×2 factorial design with conditions (irrelevant *v.* no irrelevant cues in the initial task) and transfer tasks (same location relevant *v.* different location relevant) as the two independent variables.

Method

Subjects. There were 40 children as subjects in this experiment: these were pupils at two schools in south London. Only one age level was considered. The subjects' mean age was 5 years 4 months (range 5 years to 5 years 7 months). The different groups were equated in terms of mean scores on the Peabody Pictorial Vocabulary Test (PPVT).

Procedure. Each subject learned the two discrimination tasks in one or two sessions. Through both tasks the subject and experimenter sat on opposite sides of the table. Before the first task the experimenter said: 'Now I'm going to show you two cards a lot of times, and each time I do it I'm going to hide a sweet behind one of the cards. You've got to learn how to find the sweet each time.' The experimenter would then begin the first task. In each trial the two cards were presented vertically, each being stuck into a narrow slot in a block of wood, which was 2 in. high and 8 in. wide. The sweet was always hidden on the block which carried the correct card, and was placed just behind that card. In all cases the cards were rectangular, 8 in. high and 4 in. wide, and white.

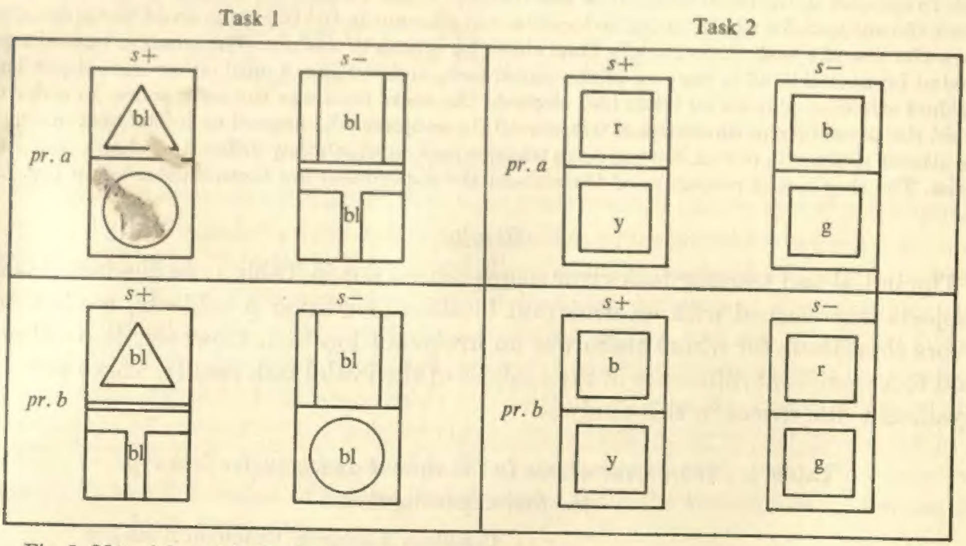


Fig. 1. Material used in the first experiment. bl = black; r = red; g = green; b = blue; y = yellow.

The distinctive feature of the material used in this experiment, some of which is illustrated in Fig. 1, was that these cards were clearly divided by a straight line into a top and a bottom half, and that in each half there was a coloured shape. In all tasks the cues which were relevant to the discrimination came from one location only. For example, in task 1 (in Fig. 1) there were two pairs of cards, each pair being presented on half the trials. In this task the stimuli on the cards were all black but differed in shape. The correct card always contained a triangle and the incorrect card a vertical rectangle; thus the triangle and rectangle were the shapes which distinguished whether or not the card was correct.

Note that these two shapes always appeared in the top half of the cards. The shapes which appear in the bottom half of the stimuli do not signify whether or not the card is correct. Thus there is both a relevant location, in which the cues indicate whether or not the card is correct, and an irrelevant location in which the cues do not carry any information about the correctness of the card. Similarly in task 2 there is also a relevant and an irrelevant location. In this task the stimuli are all the same shape and differ only in colour. The correct card always contains the yellow and the incorrect card the green square, and these are always in the bottom half, which is, therefore, the relevant location. The colours in the top half, which is the irrelevant location, do not indicate whether or not the card is correct. The position of the correct card was varied in a Gellerman series (Gellerman, 1933). In tasks in which two pairs of cards were used, the order of presentation of these pairs was also varied in a Gellerman series.

Twenty children learned an *initial task* in which the material was characterized by both a relevant and an irrelevant location (condition A). Both tasks illustrated in Fig. 1 are of this type. The other 20 subjects learned an initial task in which there was only a relevant location and no irrelevant location (condition B). In such tasks only one half of the cards contained stimuli, while the other half was completely blank: thus, if the top half was relevant, both cards in each trial would contain different stimuli in the top half, and no stimuli in the bottom half. This meant that with the latter kind of task only one pair of stimuli was used. In both types of initial task learning was to a criterion of nine out of 10 successive correct trials. Half the children in each condition learned a task in which the stimuli on the cards differed in shape, the other half learned one in which the stimuli differed in colour. Also half the children learned this task with the top location relevant, and other half with the bottom location relevant.

Half of the children in each condition then went on to a *transfer task* in which the same location was relevant as in the initial task. The other half went on to one in which the other location was relevant and the previously relevant location was now irrelevant. Comparison of these two types of task was taken to be the definitive test of whether or not the subjects continued in the transfer task to respond to the same location as was relevant in the initial task. If such transfer did take place, the subjects for whom the same location was relevant in the two tasks could be expected to learn the transfer task more rapidly than those for whom it was not. The transfer task always started immediately after the end of the initial task, and continued until either the subject had reached criterion or until 50 trials had elapsed. The score used was the error score. In order to avoid the possibility of dimensional transfer all the subjects who learned an initial task in which the stimuli differed in colour, were given a transfer task in which they differed in shape, and vice versa. The shapes and colours used throughout the experiment are those illustrated in Fig. 1.

Results

The initial and transfer task error scores are set out in Table 1. In the initial task subjects who learned with an irrelevant location (condition A subjects) made more errors than those for whom there was no irrelevant location. However, this did not lead to a significant difference in the analysis of the initial task results. There were no significant differences in this analysis.

Table 1. *Mean error scores in the initial and transfer tasks of the first experiment*

Transfer groups	Condition A subjects (irrelevant cues in IT) Mean	s.d.	Condition B subjects (no irrelevant cues in IT) Mean	s.d.
		Initial task		
Same location in both tasks	9·9	3·3	11·7	6·0
Different location in each task	9·6	4·9	13·4	6·7
		Transfer task		
Same location in both tasks	6·7	3·1	9·2	1·9
Different location in each task	11·4	4·5	10·1	2·8

In the transfer task the subjects who had learned under condition A showed considerable locational transfer. Those of them for whom the same position was relevant in both tasks learned the transfer task a great deal more rapidly than those for whom it was not. The condition B subjects showed the same pattern but to a much smaller extent.

The analysis of variance of the transfer test scores, which is presented in Table 2, produced a significant conditions × transfer task ($P < 0{\cdot}05$) interaction. Tukey tests showed that, within condition A, there was a significant difference between the

subjects for whom the same location was relevant as previously and those for whom it was not ($P < 0.05$). However, within condition B, this difference was not significant. This indicates that the interaction was caused by the fact that there was locational transfer after a task in which there had been an irrelevant as well as a relevant location (condition A), but not after one in which there was only a relevant location (condition B).

These results therefore demonstrate that locational transfer is possible in young children. They also suggest that this transfer is dependent on the existence of an irrelevant location as well as a relevant location in the initial task.

Table 2. *Analysis of the transfer task scores in the first experiment*

Source	M.S.	D.F.	F	P
Condition (C)	78·400	1	0·572	n.s.
Transfer task (T)	3·600	1	11·480	0·01
C×T	36·100	1	5·286	0·05
Residual error	6·829	76	—	—

Experiment II

The first experiment established that young children are able to transfer a response from one task to another on the basis of the location which provided the relevant cues in the initial task. Usually in traditional discrimination tasks the location and the dimension of the relevant and the irrelevant cues are confounded. For example, in a colour and shape task, colour cues come from the centre and shape cues from the contour. It follows that what has been interpreted as dimensional transfer may in fact be locational transfer of the kind which was demonstrated in the first experiment. In other words, if, for example, a subject continues to respond to colour differences in a transfer task, after learning an initial task in which colour was relevant, this may be transfer to the dimension of colour, but it may equally be transfer of a response to cues coming from the centre.

If the question of the relevance of locational to dimensional transfer is to be pursued, the next step must involve an experimental situation where dimensional transfer is possible, but where location and dimension are no longer confounded. This was the aim of the second experiment, in which the location and dimension of relevant cues were independently varied. In this experiment, all the subjects learned two tasks in which there were both a relevant and an irrelevant dimension and also a relevant and irrelevant location. With half the subjects for whom the same dimension was relevant in both tasks, the same location was also relevant in both tasks. However, with the other half the location of the relevant cues changed between tasks. Similarly with half the subjects for whom the relevant dimension differed between tasks the same location was relevant in both tasks, while with the other half the relevant location as well as the relevant dimension differed between tasks.

This design made it possible to investigate the relative strengths of dimensional and locational transfer with this kind of material. Two age levels (5- and 7-year-olds) were considered in this experiment.

Method

Subjects. The subjects in this experiment were 160 children, who came from three south London schools. Two age-groups were considered: one with a mean age of 6 years 11 months (range 6 years 6 months to 7 years 3 months), the other with a mean age of 5 years 1 month (range 4 years 10 months to 5 years 5 months). There were 80 children in each age-group. The different subgroups were matched in terms of mean scores on the PPVT within each age-group.

Procedure. The location and the dimension of the relevant cues were independently varied by using the same kind of material as in the previous experiment, but changing it to the extent that in any one task the cues were characterized by two variable dimensions as well as by two locations. This is illustrated in Fig. 2, which shows the material used in two tasks.

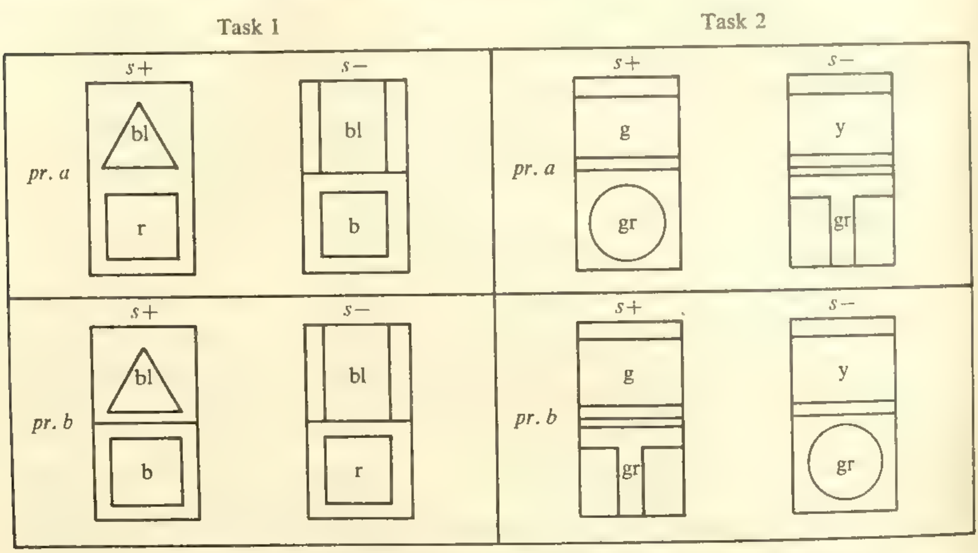


Fig. 2. Material used in the second experiment. b = blue; bl = black; y = yellow; r = red; gr = grey; g = green.

In both tasks, and indeed in all cases in this experiment, there was not only a relevant and an irrelevant location, but also a relevant and an irrelevant dimension. In task 1 (in Fig. 2) the relevant location is the top half and the irrelevant location the bottom. It is the cues in the top half only which indicate whether or not the card is correct. In this task also the relevant dimension is shape and the irrelevant dimension colour, since the relevant cues differ from each other in shape only, while the irrelevant cues differ from each other in colour only. To put it another way, shape is the relevant dimension because in order to respond correctly the subject has to remember the shape of the relevant cue. Remembering its colour will not solve the discrimination. In task 2 it can be seen that the relevant location is again the top half, but that the relevant dimension is colour, since the relevant cues differ from each other in colour. The irrelevant location is the bottom half and the irrelevant dimension shape, since the cues in the bottom half are irrelevant to the solution of the discrimination, and these cues are the ones that differ in shape. This task cannot be learned on the basis of shape differences.

With tasks such as these it is possible to measure *locational transfer* by comparing performance in transfer tasks in which the relevant cues are in the same location as in the initial task with performance in transfer tasks in which the other location contains the relevant cue. If the former are learned more rapidly than the latter it can be concluded that the subjects do transfer information about the location of the relevant cue. Similarly it is possible to measure *dimensional transfer* by comparing performance in transfer tasks in which the same dimension is relevant as previously with those in which the previously irrelevant dimension is made relevant (e.g. where colour differences are relevant in the initial task and shape differences in the transfer task). This makes i

possible to have four kinds of transfer task, which are as follows: (1) same dimension and same location relevant as in initial task; (2) same dimension relevant, but different location relevant; (3) different dimension relevant but same location relevant; (4) different dimension and different location relevant. Subjects who transfer information only about dimensions should learn tasks 1 and 2 with equal ease, and both of these more easily than tasks 3 and 4. Subjects who transfer information only about location should learn tasks 1 and 3 with equal ease and these much more easily than tasks 2 and 4. This design is schematized in Table 3.

These were the main comparisons of this experiment. The design was a $2 \times 2 \times 2 \times 2$ factorial one, with 10 subjects in each cell. The independent variables were age, dimensional transfer (tasks with same dimension relevant as previously compared with tasks with the other dimension relevant), locational transfer (tasks with same location relevant compared with those in which the other location is relevant) and material (tasks in which colour was relevant compared with those in which shape was relevant). This last variable was included not because it raised any important question, but because it had been found in pilot studies that shape discriminations tended to be learned more rapidly than colour discriminations. All the subjects learned the first task to a criterion of nine out of 10 correct successive trials. The transfer task continued until either this criterion was reached or until 50 trials had elapsed.

In each task the position of the correct cards and the order of the pairs of cards were varied in the same way as in the first experiment.

Table 3. *Schema of the design of the second experiment*

(Same = same in both tasks. Different = different in each task.)

	Task 1 subjects	Task 2 subjects	Task 3 subjects	Task 4 subjects
Dimension	Same	Same	Different	Different
Location	Same	Different	Same	Different

Results

Initial task. The initial task mean error scores are presented in Table 4. The only significant difference in the analysis of these scores was between ages ($F = 53{\cdot}490$; d.f. $= 1, 145$; $P < 0{\cdot}001$). This was due to the fact that the older group learned more rapidly than the younger. Since there were no significant differences between sub-groups within the two age levels, it can be assumed that differences in the transfer task scores are due to variations in the experimental conditions.

Transfer task. The transfer task error scores are also presented in Table 4. The scores were analysed together in one total analysis, and in addition the scores of the two age-groups were analysed separately. Three main questions are raised by this experiment. The first is whether there is dimensional transfer when locational influences are controlled. The second is whether there is locational transfer when dimensional effects are controlled. The third is whether there is any interaction between age and the strength of either of these kinds of transfer.

The answer to the first question is that there is evidence for dimensional transfer at both age levels. The combined analysis produced a significant dimension main term ($F = 22{\cdot}427$; d.f. $= 1, 145$; $P < 0{\cdot}001$). There were also significant dimension terms in both separate analyses ($F = 25{\cdot}950$; d.f. $= 1, 72$; $P < 0{\cdot}001$ in the case of the older group, and $F = 4{\cdot}069$; d.f. $= 1, 72$; $P < 0{\cdot}05$ in the case of the younger group).

The answer to the second question is that there is also evidence for locational transfer, but only in the case of the younger group. The combined analysis did

produce a significant location main term ($F = 18{\cdot}983$; d.f. = 1, 145; $P < 0{\cdot}001$), but it also produced an age × location interaction ($F = 4{\cdot}943$; d.f. = 1, 145; $P < 0{\cdot}05$). Moreover, in the separate analyses the location term was significant only in the case of the younger group ($F = 17{\cdot}164$; d.f. = 1, 72; $P < 0{\cdot}01$), and not in the case of the older group ($F = 3{\cdot}001$; d.f. = 1, 72; P n.s.). Younger children transfer a locational response very strongly in this situation; older children do not.

Table 4. *Mean error scores in the initial and transfer tasks of the second experiment*

(Task 1: same dimension and same location relevant in IT and TT. Task 2: same dimension, different location relevant in IT and TT. Task 3: different dimension, same location relevant in IT and TT. Task 4: different dimension and different location relevant in IT and TT.)

	Task 1 subjects		Task 2 subjects		Task 3 subjects		Task 4 subjects	
Dimension relevant...	Shape	Colour	Shape	Colour	Shape	Colour	Shape	Colour
				Initial task				
5-year group								
Mean	10·8	12·5	10·5	13·1	10·1	12·1	11·8	10·9
S.D.	5·5	5·4	5·3	6·1	5·7	6·7	4·6	6·5
7-year group								
Mean	6·5	7·6	5·6	5·5	5·6	4·8	4·5	7·0
S.D.	3·3	3·5	4·0	3·5	4·2	3·5	2·6	4·5
				Transfer task				
5-year group								
Mean	5·1	6·6	8·3	10·2	5·7	7·9	12·0	13·9
S.D.	3·6	4·9	5·2	5·5	3·7	3·7	6·3	5·5
7-year group								
Mean	2·1	3·4	2·7	4·1	5·2	7·6	8·4	9·3
S.D.	2·1	2·8	2·8	3·5	2·9	3·9	4·9	5·9

The answer to the third question is that two differences were noted between age-groups. The first is that older children made fewer errors than younger children, which produced a significant ages main term ($F = 21{\cdot}459$; d.f. = 1, 145; $P < 0{\cdot}001$). The second is the result noted above, which was that, while younger children showed very strong locational transfer, older children did not. Since the combined analysis produced no other significant interactions with age, it has to be assumed that there is no evidence for a change in the strength of the dimensional response with age.

The only other significant result in the three analyses was a significant materials term in all three analyses ($F = 5{\cdot}405$; d.f. = 1, 145; $P < 0{\cdot}025$ in the combined analysis), which showed that the shape tasks were learned with fewer errors than the colour tasks.

Inspection of the mean scores in Table 4 bears these conclusions out. The older group consistently made less errors in the two transfer tasks, in which the same dimension was relevant as in the initial task, than in the transfer tasks in which the other dimension was relevant. Although in this age-group there does appear to be some superiority in tasks in which the same location is relevant as before over those in which it is not (less errors in task 1 than in task 2 and in task 3 than in task 4), this effect must be disregarded since it falls short of significance.

In the younger group the most marked differences were between tasks 1 and 3 on the one hand and tasks 2 and 4 on the other. The former tasks were those in which the same location was relevant as in the initial task while the latter were those in which the other location was relevant. This indicates that the younger subjects tended very strongly to respond on the basis of the location of the previously relevant cues, despite changes in the dimensions distinguishing the cues in this location. There was also some evidence of dimensional transfer in this group (task 1 better than task 3 and task 2 better than task 4), and as reported above this tendency was significant. However, the pattern of these scores as well as the significance levels of the analysis demonstrates that this group tended to transfer on the basis of location a great deal more than on the basis of dimension.

Thus the conclusions that can be drawn from these results are as follows. (1) Both age-groups showed dimensional transfer, both when the location of the relevant cue was changed and when it was held constant between tasks. (2) The younger group also showed considerable locational transfer, both when the relevant dimension was changed and when it was held constant between tasks. (3) Older subjects made fewer errors than younger subjects. (4) Shape tasks were learned with fewer errors than colour tasks.

Discussion

The experiments produced three main results. The first was that young children do transfer from one task to another a response to the location of relevant cues. The second was that 5-year-old children show very strong locational transfer in a task in which dimensional transfer is also possible. The third was that both 5- and 7-year-olds did show dimensional transfer when locational effects were controlled. These results have two kinds of implication. The first is methodological, and concerns the use of discrimination learning techniques. The second is theoretical, and concerns the importance of the locational response for the study of perceptual development and of pattern recognition.

As far as discrimination learning methods are concerned, it is now possible to suggest that, where dimensional transfer has been reported in young children, a considerable proportion of it may be due to locational transfer. Young children do transfer a locational response, even when dimensional transfer is also possible. Since dimension and location are typically confounded in discrimination learning studies, it is possible that transfer effects are equally confounded in these studies.

It is, of course, true that in the two experiments reported here the locations were explicitly and quite grossly separated from each other. It may be that locational responses would not be as strong with the material used in traditional discrimination tasks. Nevertheless, since such a large locational response has been found in young children, there is a strong possibility that this will affect their transfer performance in any task in which different cues come from different locations. This is an important conclusion, since it is consistently the case that studies which report dimensional transfer in young children use material in which the cues from the different dimensions come from clearly separable areas. For example, all such experiments reviewed by Kendler & Kendler (1962), Tighe & Tighe (1966), Shepp & Turrisi (1967) and Wolff (1967) in which dimensional transfer is found in young children use one of three

combinations of dimensions: colour and shape, colour and size, and colour and orientation. None uses combinations such as colour and texture, or shape and orientation, where cues from different dimensions come from the same general location. Developmental studies of transfer in discrimination learning may therefore raise more complex questions than had been suggested hitherto. Future studies in this area should either try to control for these complexities, or should look for more appropriate methods for investigating transfer in young children. Categorization (Clarke *et al.*, 1967) or patterning (Frith, 1968) studies might be more suitable.

However, the finding that young children record and respond on the basis of the location of a cue within a display, even when the display is moved from trial to trial, has theoretical implications as well. Bryant (1969) has shown that one method which young children use in remembering the orientation of lines is to record the absolute location of lines within the visual display. These subjects remember the direction of oblique lines if the location of these lines is kept constant between presentation and the recognition test. However, they perform at chance level if the location of the lines within the stimulus card differs between presentation and the recognition test. This suggests that children utilize information about location of cues in tasks which involve information about orientation, and therefore that this response may have quite wide implications in the study of their behaviour.

The locational response is also relevant to studies of pattern recognition in young children. Patterns in which the location of distinctive features is constant should be more easily learned than those in which it varies. This suggestion is supported by a recent finding (Bryant & Weightman, 1969) that severely subnormal children, who also have a strong tendency to transfer locationally (Bryant, 1967*b*), do make more errors in a pattern recognition task when the location of the distinctive features varies than when it is constant. A locational strategy is obviously an inappropriate one if the child is to learn to read, since the location of the features which distinguish alphabetical letters varies considerably both between and within letters (Bryant & Weightman, 1969). It is possible, therefore, that the reason that the locational strategy is found in children of 5 but not in children of 7 is that the latter group have to extinguish this strategy as a result of their experience in learning to read.

References

Bryant, P. E. (1967*a*). The role of the irrelevant stimulus dimension in transfer by young children. *Proc. Symposium on Discrimination Learning, University of Sussex.*

Bryant, P. E. (1967*b*). Selective attention and learning in severely subnormals. *Proc. 1st Int. Congr. on Scientific Study of Mental Deficiency, Montpelier.*

Bryant, P. E. (1969). Perception and memory of the orientation of visually presented stimuli by young children. *Nature, Lond.* **224**, 1132–1133.

Bryant, P. E. & Weightman, J. (1969). Discrimination learning and the learning of letters by severely subnormal children. *J. ment. Defic. Res.* **13**, 221–234.

Clarke, A. M., Cooper, E. M. & Clarke, A. D. B. (1967). Task complexity and transfer in the development of cognitive structures. *J. exp. Child Psychol.* **5**, 562–576.

Ehrenfreund, D. (1948). An experimental test of the continuity theory of discrimination. *J. comp. Psychol.* **41**, 408–422.

Frith, U. (1968). Feature extraction in young children. *Bull. Br. psychol. Soc.* **22**, 49.

Gellerman, L. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.* **42**, 206–208.

House, B. J. & Zeaman, D. (1962). Reversal and nonreversal shifts in discrimination learning of retardates. *J. exp. Psychol.* **63**, 444–451.

Kendler, H. H., Glucksberg, S. & Keston, R. (1961). Perception and mediation in concept learning. *J. exp. Psychol.* **61**, 186–191.

Kendler, H. H. & Kendler, T. S. (1962). Horizontal and vertical processes in problem solving. *Psychol. Rev.* **69**, 1–16.

Shepp, B. E. & Turrisi, F. D. (1967). Learning and transfer of mediating responses in discrimination learning. In N. R. Ellis (ed.), *International Review of Research in Mental Retardation*, vol. 2. New York: Academic Press.

Tighe, T. J. & Tighe, L. S. (1966). Discrimination learning: two views in historical perspective. *Psychol. Bull.* **66**, 353–370.

Tighe, T. J. & Tighe, L. S. (1968). Differentiation theory and concept shift behaviour. *Psychol. Bull.* **70**, 756–761.

Turrisi, F. D., Shepp, B. E. & Eimas, P. D. (1969). Inter- and extra-dimensional shifts with constant and variable irrelevant dimensions in the rat. *Psychon. Sci.* **14**, 19–20.

Wolff, J. L. (1967). Concept shift and discrimination reversal in humans. *Psychol. Bull.* **68**, 369–408.

Zeaman, D. & House, B. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis (ed.), *Handbook of Mental Deficiency*. New York: McGraw-Hill.

(*Manuscript received* 23 *October* 1969; *revised manuscript received* 13 *January* 1970)

Br. J. Psychol. (1971), 62, 1, pp. 13-20
Printed in Great Britain

CHILDREN'S JUDGEMENTS OF DURATION

By B. M. HERMELIN and N. O'CONNOR
MRC Developmental Psychology Unit, Gordon Street, London

[illegible] heard one after the other.

Experiment I

In a previous pilot experiment we had found that 5-year-old normal children could not learn to discriminate between stimuli of different durations. Two questions arose from this: was the difficulty of temporal judgements for young children one of orientation or one of perception, and did the modality of the stimuli and their manner of presentation affect such judgements? The following two experiments are concerned with these questions.

There are relatively few experiments in the literature which investigate the ability of children to match, or to discriminate between, stimuli of different durations. In those experiments which have been reported, such as those of Fraisse & Orsini (1958) or Smith & Goldstone (1957), different paradigms have been used, some concerning the interval between signals, and others requiring judgement of the duration of stimulation. The results are contradictory, and Fraisse states that there seems to be 'no intuition of duration as such' in young children.

This is in accordance with Piaget's (1969) view that the young child is unable to judge duration independently of speed, distance or some similar variable. In the terms of the two-stage discrimination model put forward by Sutherland (1959) the child seems unable to 'switch in' the appropriate temporal analysers. However, it is also possible that it is the second stage in the discrimination process, i.e. the perceptual analysis of information, which may be deficient. The child may fail, therefore, either because of the inefficiency of these 'switching' mechanisms, or because he is unable to estimate durations with accuracy. In the first experiment to be reported, we investigated the effect of specific instructions on the switching of attention to the appropriate temporal dimension.

Method

Subjects. Children aged between 5 years 1 month and 5 years 11 months were given a pretest to ascertain whether they understood the meaning of the words 'same' and 'different', and could use them appropriately. Each child was individually presented with 12 3½ × 3½ in. cards on which two 1 × 1 in. white squares were pasted. These squares showed sample drawings from Gibson's *et al.* (1962) letter-like forms, varying in degree of completeness. Six of the cards showed a pair of identical figures, while the other six showed one complete and one incomplete shape. The child was asked to say whether the drawings on any one card were 'the same' or 'different'.

Only children who gave an errorless performance on this task were included in the main experiment. After 40 children had been selected, they were divided into two matched groups with a mean CA of 5 years 5 months each.

Apparatus. This consisted of a box containing three timers connected in cascade, and two rotary switches for selecting duration of stimuli. The stimuli were displayed visually on a 12 × 12 cm opaque screen. They were two edgelit perspex lines, 2 cm wide and 10 cm long, with a gap of 5 cm between top and bottom lines. The lines were formed by illuminating one edge of a perspex sheet by means of a small lamp. The edge opposite the lamp was cut to the size required, and pressed against a translucent screen. The opposite side of the screen was viewed by the subject. The time lag between the disappearance of the first, top line, and the appearance of the second, bottom line, was 0·5 sec. Luminosity of each line was 50 foot lamberts.

There were four possible settings, shown below:

Setting	First stimulus (sec.)	Interval (sec.)	Second stimulus (sec.)
I	6·0	0·5	6·0
II	2·0	0·5	2·0
III	6·0	0·5	2·0
IV	2·0	0·5	6·0

Procedure. Each child was individually tested for 24 trials with the same randomized series, containing six instances of each of the four conditions given above. After a pair of lines had been presented, the child was asked to say whether the two had been 'the same' or 'different'. The experimenter reinforced correct responses by saying 'yes' and incorrect ones were discouraged by saying 'no'. The two groups were given the following instructions:

Group A: specific instructions. You will see a light come on here, and then it will go off. Then you will see a light come on there and then it will go off too. I want you to watch the lines carefully and tell me if they stay on for the same time. If one stays on for the same time as the other one, I want you to say 'same'. But if one line stays on for a longer time than the other line, I want you to say 'different'.

Group B: general instructions. You will see a light come on here and then it will go off. Then you will see a light there and then it will go off too. I want you to watch the lines very carefully and tell me if you think they were the same or different.

Table 1. *Error scores, means and standard deviations for two groups under two conditions of instructions*

(Maximum scores = 6 for each stimulus arrangement.)

	Stimulus arrangements (sec.)			
	2–2	6–6	2–6	6–2
Specific instruction group (n = 20)				
Mean	2·40	2·10	1·85	2·55
S.D.	1·50	0·97	1·18	1·19
General instruction group (n = 20)				
Mean	2·60	2·30	3·65	3·70
S.D.	1·31	1·38	1·23	1·17

Results

Results were first scored according to the number of incorrect responses in each of the two conditions. The mean error scores are shown in Table 1. An analysis of variance of these scores resulted in significant groups (F = 32·42; d.f. = 1, 38; P = 0·001) and conditions (F = 4·40; d.f. = 1, 38; P = 0·05) effects, as well as in a

significant interaction ($F = 8 \cdot 12$; d.f. = 1, 76; $P = 0 \cdot 01$). This interaction demonstrated that when the two lights were on for the same time, the general and specific instruction groups performed similarly. However, with pairs of stimuli of different duration, those receiving specific instructions did better than those who did not. When the two illuminated lines were of different durations, a long duration stimulus of 6 sec. could be followed by a short one of 2 sec. or vice versa. An analysis carried out for these conditions yielded a significant interaction ($F = 9 \cdot 20$; d.f. = 1, 76; $P = 0 \cdot 01$), which showed that there was no difference between them for the group receiving only general instructions. However, those children who had been told that duration was the relevant discriminable variable made significantly fewer mistakes when a short duration stimulus was followed by a long one than when the longer was presented first.

Though those children who received specific instructions orientating them towards the time dimension did better than those not so 'set', this effect was not sufficient to result in a performance of 8 out of 10 correct responses over a series of 24 trials. Five out of 20 in the specific, and 1 out of 20 in the general instruction group reached this criterion. The differences of these frequencies as tested by χ_2 was not significant.

Discussion

Two main points for discussion arise from these results. They are firstly the effect of giving children a specific set towards a temporal dimension. Secondly, there is the accidental finding that, when such a set is given, a short stimulus followed by a long one is perceived more easily as a difference in duration than is a long followed by a short.

The experiment was planned to test possible explanations for a previous result that young children had difficulty with temporal discriminations. This might be due to at least one of two factors. Either the child can really not discriminate between different durations, and/or the appropriate analyser has not been switched in, and the stimuli are thus judged according to some other, in this instance, irrelevant dimension, such as extension in space, luminosity etc. Our results indicate that both these factors operate. Children who were given a set towards judging according to duration did better than those not so instructed. Thus specific instructions resulted in a switching in of the appropriate temporal analysers. However, this set was not sufficiently effective to enable the subjects to reach a criterion of consistently correct responses. Thus, though these children were presumably looking for time differences, they could not always judge them correctly. That they were, in fact, attempting to judge duration is indicated by the distribution of errors, which was similar for trials in which the lights had the same durations and those in which they had different durations. This is in contrast with the results for the group who were not specifically set towards the time dimension. As these children gave significantly more correct responses when the lights lasted for the same duration than when they lasted for different durations, one must conclude that the predominant response of 'same' was directed towards a non-temporal dimension. Both illuminated lines were, for instance, of the same colour, luminosity and extension in space. Thus in a multidimensional classification task they were indeed 'the same' in more ways than they were 'different'. It seems from the results that specific instructions resulted in appropriate orientation,

while in the absence of such a specific instruction orientating the child towards the relevent dimension, the temporal analysers were not readily switched in.

An additional finding was that a short followed by a long duration stimulus was easier to judge as being different than a long followed by a short one. One indication for an explanation of this phenomenon came from a qualitative observation of the responses. Children who were correct often tended to give a response of 'different' as soon as the second stimulus had been on for more than 2 sec. rather than wait for the termination of the light for a possible full 6 sec. duration. Thus while judgements of short followed by long stimuli could be made soon after about 4·5 sec. of total display time, such judgements had to be deferred for about 8·5 sec. in conditions in which the long stimulus was followed by the shorter one. Thus the total amount of display time elapsing before judgements became possible may be relevant for these results.

This result is in contrast with those reported by Huppert & Singer (1967) for judgements of auditory durations. They found a positive temporal after-effect, i.e. the duration of a signal following an interpolated tone was underestimated relative to the pretest duration. Our results indicate a negative time error, where the second of two stimuli tends to be over-estimated in relation to the first, thus in our case reducing the difference between a long followed by a short signal. With adults such negative time errors have usually been found only with longer time intervals between the two stimuli. However, it is at least possible that in young children negative time errors are produced over shorter interstimulus intervals.

Experiment II

The first experiment used successively presented visual stimuli in a temporal discrimination task. The relative failure of the subjects to judge two durations as either being the same or different may have been specifically related to the stimulus modality used, or to the successiveness of the exposures. Thus stimuli in another sense modality, or a simultaneous rather than a successive stimulus exposure, might yield different results.

It has been argued that hearing is a process which integrates successive stimuli. Although Savin (1967) suggests that the results of many experiments tend to confirm such a view, simultaneous sounds are also obviously integrated by the auditory system, as, for example, in the appreciation of harmony or the recognition of different speech accents. Conversely, though the eye deals with a wide visual field which presents information simultaneously, successive focusing on different points also occurs. The experiment reported below was designed to investigate the effect of simultaneous and successive presentation of stimuli on temporal and non-temporal discriminations when stimuli were presented in either the visual or auditory modality.

Method

Subjects. The same pretest procedure as in Expt. I was used to select 80 subjects aged between 5 years 1 month and 5 years 11 months. They were divided into eight matched groups of 10 subjects on the basis of age. The average age of each group was 5 years 4 months.

Apparatus. This consisted of the opaque glass screen used in the first experiment, earphones, and boxes containing timing and control devices as in the first experiment. Visual as well as auditory sets of stimuli were presented. Lights had an intensity of 50 foot lamberts and tones a

level of 70 db. Illuminated lines extended for either 5, 10 or 15 cm. They could also vary in exposure duration of either 2, 4 or 6 sec.

The auditory stimuli were tones which varied in pitch and in duration. They could be high, medium or low pitch, i.e. 1200, 800 or 400 Hz. Their duration, as the duration of the lights, could be either 2, 4 or 6 sec. Each member of a pair of auditory stimuli was presented to one ear through the earphones. Which stimuli were directed to the right and left ear respectively was randomly varied according to a balanced design.

Procedure. The procedure was designed to compare the effect of stimuli presented either simultaneously or successively, in either the visual or auditory modalities, and in terms of the dimensions of space or pitch, as compared with time. In all conditions 24 pairs of stimuli were presented to the subjects who had to judge whether these stimuli were the same or different. Each correct response was reinforced by the experimenter saying 'good' and each incorrect one was followed by the experimenter saying 'no'. The pairs of stimuli were either visual or auditory and varied in one of two ways. Either they varied in duration or alternatively they varied in another dimension, e.g. length in the case of visual and pitch in the case of auditory stimuli. When duration was invariant, this second dimension varied, and when the alternative dimension was invariant, duration was varied from trial to trial. There were eight combinations of variables, which are set out in Table 2.

Table 2. *Presentation of stimuli in Expt. II*

Groups	Conditions	Presentation	Modality
I	Duration variable, length invariant	Successive	Visual
II	Duration invariant, length variable		
III	Duration variable, length invariant	Simultaneous	Visual
IV	Duration invariant, length variable		
V	Duration variable, pitch invariant	Successive	Auditory
VI	Duration invariant, pitch variable		
VII	Duration variable, pitch invariant	Simultaneous	Auditory
VIII	Duration invariant, pitch variable		

Subjects were divided into eight groups of 10 children each. Each group was tested under one of the conditions shown in Table 2. The 24 pairs of stimuli presented were randomized in such a way that equal and unequal pairs, and first or second exposure of one or the other signal of a stimulus pair, were balanced over the series. When duration was invariant, exposure time was always 4 sec., but when it was variable, stimuli were either both exposed for 2 sec. or both for 6 sec., or one for 2 sec. and one for 6 sec. A similar rule applied to the constancy or variation of length and pitch. When duration was varied an intermediate length (10 cm) or pitch (800 Hz) was used. When the non-temporal dimensions were variable, combinations of the extremes (5 and 15 cm; 400 and 1200 Hz) were used, while duration was held constant at 4 sec.

Instructions were always similar, e.g. 'Two lights (tones) will come up here and there. Watch (listen to) them carefully and when they have gone, tell me whether they were the same or different.' Successive displays were separated by a 0·5 sec. gap. Simultaneous displays always commenced at the same time, and the visual stimuli appeared one above the other, with a 5 cm gap between them.

Results

The results were scored in terms of number of errors for each child under any of the eight conditions. These mean error scores are illustrated in Fig. 1.

Testing for simple effects, by means of Sandler's A test (1955), it was found that there were more errors in durational than in non-durational judgements ($A = 0{\cdot}0736$; $P = 0{\cdot}001$). As no specific instructions about the dimension to be judged had been given in this experiment, this might have been expected from the previous results. In judgements of duration, successive presentation of stimuli resulted in more errors than simultaneous exposure ($A = 0{\cdot}0943$; $P = 0{\cdot}001$). In non-temporal judgements

the reverse was the case in the auditory modality ($A = 0{\cdot}1332$; $P = 0{\cdot}001$) i.e. more errors were made with simultaneously than successively presented sounds.

With duration judgements, more errors occurred when the stimuli were lights than when they were tones ($A = 0{\cdot}2330$; $P = 0{\cdot}05$). Conversely, in non-temporal conditions more errors were made with auditory than with visual stimuli ($A = 0{\cdot}0860$; $P = 0{\cdot}001$).

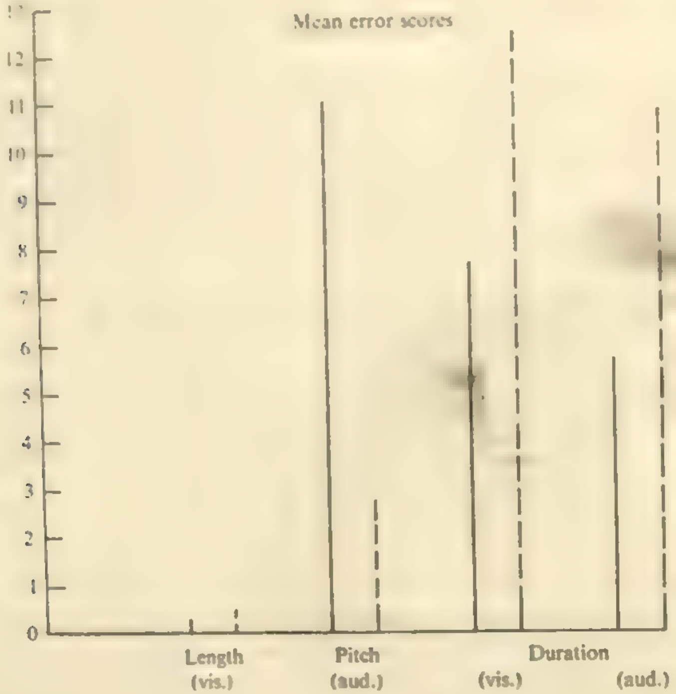


Fig. 1. —, simultaneous; – – –, successive.

Table 3. *Error scores for eight conditions*

(Total possible errors = 24 for each condition.)

	Non-durational		Durational	
	Visual	Auditory	Visual	Auditory
Simultaneous				
Mean	0·20	11·00	7·70	5·70
S.D.	0·42	3·70	5·68	4·06
Successive				
Mean	0·50	2·80	12·50	10·80
S.D.	0·71	4·02	0·97	1·62

An analysis of variance was based on differences in error scores between durational and non-durational conditions. The mean error scores are shown in Table 3.

The analyses yielded two highly significant main effects. Thus differences between error scores of temporal and non-temporal judgements were greater when the stimuli were successively than when they were simultaneously presented ($F = 44{\cdot}37$;

d.f. = 1, 36, $P < 0.001$). Also, these differences are greater with visual than with auditory stimuli ($F = 41.72$, d.f. = 1, 36, $P < 0.001$). The analysis of the temporal versus non-temporal difference error scores also showed a significant interaction (F = [illegible], d.f. = 1, 36, $P < 0.01$). This is accounted for by the large difference between difference error scores in auditory as compared with visual simultaneous conditions. It also takes into account the considerable difference between simultaneous and successive auditory scores. For an interpretation of these differences, we must refer to the original error scores from which these difference scores were derived. These show that whereas in three of the four main conditions errors for durational judgements were greater than for non-durational ones, the reverse is the case in conditions of simultaneously presented auditory stimuli. In these conditions there are fewer incorrect durational than non-durational judgements.

Discussion

The results must be discussed in terms of three paradigms: the relative difficulty of durational as compared with non-durational judgements, and the effects on these judgements of manner (i.e. simultaneous or successive) and modality (i.e. visual or auditory) of stimulus presentation. On the whole, durational judgements were more difficult than the non-durational ones, though this difference was more marked when lights rather than when sounds were the stimuli.

Durational judgements were more efficient when the stimuli were exposed together than when one followed the other, and this was true with lights as well as with sounds. The patterns of responses for durational judgements are very similar for both stimulus modalities, indicating similar response strategies. In contrast to one of the previously advanced hypotheses, simultaneous presentation of stimuli varying in duration resulted in better performance with auditory as well as with visual signals. Thus in these conditions successive stimulus presentation did not result in any advantage for the auditory sense. The slight overall advantage of auditory over visual temporal conditions might have been due to the fact that the stimulus sources for the two sounds were more clearly distinct than those for the light. Both lights were exposed on the same screen, and seen by both eyes. With the sounds, one clearly came into the right and the other into the left ear.

With simultaneous stimulus presentation, where fewer errors were made when judging duration, the pair of stimuli always appeared together. They either disappeared together, or one stayed on after the other had ceased. It is thus possible that in these conditions judgements were made according to whether one or two stimuli were present at the end of a trial—rather than according to true durational estimates. Such a conclusion is supported by the results in the conditions of successive stimulus presentation. In these conditions, in which duration was probably the only cue, responses were on a chance level. An analysis of response biases showed that judgements of 'same' were made significantly more often in this than in any other condition.

In non-durational judgements of visual stimuli, where the length of pairs of lines had to be compared, performance was nearly perfect. This was independent of whether these lines were successively or simultaneously exposed. However, judgements of the pitch of auditory stimuli, although usually correct when the tones were

successively presented, were at chance level when they were sounded together. This difficulty in judging simultaneous tones may have resulted from the particular stimuli. As these were two pure tones, 1½ octaves apart (i.e. 1200 and 400 Hz), the sine waves of the higher one were contained in those of the lower. Moreover, the respective cochlea receptor fields at which microphonic potential is maximal for the sine waves generated by these two tones, lie approximately on the same axis. Thus this discrimination might have been particularly difficult. If more complex noises with different components of sine waves had been used, each tone might have preserved its distinct characteristic. Consequently no general conclusion about auditory resolution of simultaneously presented stimuli should be drawn from these results.

Our thanks are due to Miss M. L. Blane, Headmistress of Dog Kennel Hill School, Miss J. Groves, Headmistress of Heathfield Infants School, and Mr R. V. Ferris of Harvard Primary School and their staffs who helped us with this investigation.

REFERENCES

FRAISSE, P. (1963). *The Psychology of Time*. London: Harper & Row.

FRAISSE, P. & ORSINI, F. (1958). Etudes expérimentale des conduites temporelles. *Année psychol.* **58**, 1–6.

GIBSON, E. J., GIBSON, J. J., PICK, A. D. & OSSER, H. (1962). A developmental study of the discrimination of letter-like forms. *J. comp. physiol. Psychol.* **55**, 897–909.

HUPPERT, F. & SINGER, G. (1967). An aftereffect in judgement of auditory duration. *Percept. Psychophys.* **2**, 544–546.

PIAGET, J. (1969). *The Child's Conception of Time*. London: Routledge & Kegan Paul.

SANDLER, J. (1955). The *A* test of difference between the means of two correlated samples. *Br. J. Psychol.* **46**, 225–226.

SAVIN, H. B. (1967). On the successive perception of simultaneous stimuli. *Percept. Psychophys.* **2**, 479–482.

SMITH, E. J. & GOLDSTONE, S. (1957). The time sense: a normative genetic study of the development of time perception. *Percept. mot. Skills*, **7**, 49–59.

SUTHERLAND, N. S. (1959). Stimulus analysing mechanisms. In *Symposium on the Mechanization of Thought Processes*, vol 2. London: H.M.S.O.

(*Manuscript received* 14 *January* 1970)

Br. J. Psychol. (1971), 62, 1, pp. 21-26
Printed in Great Britain

CRITERIA OF SUCCESS IN A DEVELOPMENTAL STUDY OF ODDITY LEARNING

By ANN L. BROWN* and BARBARA B. LLOYD
School of African and Asian Studies, University of Sussex

Equal numbers of boys and girls in five groups of [illegible] CA and IQ were tested on a [illegible] three position planometric oddity problem. The [illegible] CA and oddity learning [illegible] examined for instrumental and verbal criteria. Four experimental conditions [illegible] the [illegible] of form (both geometric and life) and colour as the stimulus dimension for oddity [illegible]. The instrumental criterion was first met by a majority of subjects in the [illegible] year old group. Majority achievement of the verbal criterion [illegible] occurred in the [illegible] year old group. Verbal justification shifted from 'same' to 'difference' explanations with age. Initial training on either form stimuli set led to greater success than initial colour training.

This study seeks to clarify three issues concerning oddity learning in children (1) the relationship between verbal and instrumental criteria of success, (2) the relationship between CA and oddity learning in children of average intelligence, and (3) the differential difficulty of three stimulus sets colours, life forms and geometric forms.

A preliminary analysis of the criteria employed as evidence of successful solution is necessary for an understanding of the relationship of CA and oddity learning. Inhelder & Piaget (1964) used performance on an oddity problem to test knowledge of the singular class. Requiring a verbal explanation of the correct solution and reproduction of an oddity problem, they are able to report only 50 per cent of 7-year olds achieving oddity solutions. The customary practice reported in the experimental literature has been to require an instrumental criterion of six consecutive, correct responses (Gollin & Shirk, 1966) or 20 correct out of 25 (House, 1964) as evidence of oddity learning.

Instrumental criteria yield larger proportions of successful performance at lower ages than verbal criteria. Multiple criteria studies indicate that even subjects meeting an instrumental criterion of oddity learning exhibit developmental differences in the ability to verbalize the solution. Scott (1964) found that retardates failed direct verbal questions concerning oddity but were able to reach instrumental criteria on a subsequent oddity learning task. Lunzer (1968), using normal children, found a developmental difference on verbal measures, but a possible transitional age range was not explored due to his choice of 4-, 7- and 8-year-old subjects.

Results based upon instrumental criteria are inconsistent concerning the age at which oddity problems can be solved. In planometric colour oddity learning Lipsitt & Serunian (1963) reported one out of six 4-year-old subjects able to learn oddity problems, while Gollin & Shirk (1966) obtained successful performance from 42 per cent of their 24 4-year-old subjects. Samples are small and no information is furnished on IQ but studies comparing normal and retarded subjects do supply information on the relationship between MA and oddity learning. Ellis & Sloan (1959) found little

* Now at the University of Illinois. Requests for offprints should be addressed to Dr B. B. Lloyd, School of African and Asian Studies, University of Sussex.

learning in any subjects of MA less than 4 years but successful performance in a majority of subjects of MA 6 years and more. House (1964) reported failure in stereometric oddity learning using a random procedure in retardates with MAs between 3 and 5 years. However, a pictorial oddity test in which the subject is asked to 'point to the one that is not the same as the others' appears as a subtest of the Stanford-Binet at 4 years 6 months. While many studies indicate that 4-7 years is the crucial range for oddity learning in normal subjects, this study seeks precise information on the age/oddity relationship.

Finally, the use of a single stimulus dimension for all age-groups may result in the provision of tasks of differential difficulty. House (1964) has suggested that the solution to an oddity problem requires the acquisition of a chain of three responses, two observing responses and a terminal instrumental response. The first observing response is to the dimension carrying oddity, the vehicle dimension. If it is essential to first make an observing response to the relevant vehicle dimension it is necessary to take account of the developmental shift from colour to form preference influencing attention in young children (Suchman & Trabasso, 1966*a*). Several studies have shown the influence of dimensional dominance in young children's discrimination learning (Suchman & Trabasso, 1966*b*; Smiley & Weir, 1966). Adding to the colour form issue is Serpell's (1968) finding of greater form preference in young non-western children when the forms employed are silhouettes of familiar objects rather than geometric figures. Thus the third aim is to investigate the differential difficulty of colour and form as the vehicle dimension for oddity.

Method

Subjects. Restrictions on CA and IQ resulted in the selection of 80 subjects from 97 children tested. Five groups were constituted, each containing eight boys and eight girls. Subjects in groups 1 and 2 attended the same infant school, the remaining subjects being drawn from one primary school. The Stanford Binet Intelligence Scale (Third Revision), Form L–M, was administered prior to oddity testing but retained the Binet Oddity items. Table 1 presents the CA and IQ range of the five groups.

Table 1. *Descriptive data on the subjects of groups* 1–5

(CA expressed in months.)

	Groups ...	1	2	3	4	5
CA	Range	40–44	52–56	64–68	76–80	88–92
	Mean	41	54	65	78	91
IQ	Range	93–113	95–111	95–115	95–113	95–110
	Mean	108	106	109	104	103

Apparatus. A mechanical aluminium box, 30 × 12 × 10 in., was constructed with a front panel containing three 2½ in. square clear perspex windows which displayed the stimulus material. Pressure on the correct (odd) window released a brown Smartie, a pellet-shaped chocolate, through an aperture directly below the window. Between trials an additional aluminium window covered the perspex openings. An intertrial interval of about 20 sec. was necessary to reset the apparatus.

Stimulus material. Three sets of stimulus material were used: one based on colour, i.e. blue, red and yellow, one containing black geometric forms on a white field, i.e. circles, triangles and stars, and one containing black silhouettes of familiar forms, i.e. cars, boots and umbrellas.

Procedure. From the 18 possible position-counterbalanced combinations of two identical and one odd [illegible] for each material, two complete sets of [illegible] 36 randomly arranged trials [illegible] presented [illegible]. Each subject was tested on one form [illegible] and the [illegible] on successive days. A [illegible] design in which two groups were tested first on [illegible] and then [illegible] and two groups first on [illegible] followed by [illegible] the four [illegible] used: geometric form [illegible] life form [illegible] colour geometric form and colour life form.

At the beginning of each session the subject was seated facing the window panel and the experimenter [illegible] seated [illegible] behind the lid of the box. The instructions were: 'We are going to play a game. I am going to show you three windows. There is a picture in each window. The game is to find out which picture will give you a Smartie. Now if you think that this picture will give you a Smartie press it like this' (the experimenter demonstrates by pressing window 1 and is rewarded). 'If you think that this is the one which will give you a Smartie press this one' (the experimenter presses and is rewarded on 2 and 3). 'Each time I open the windows it is a new game. Only one window will give you a Smartie. See how many Smarties you can get.' Smarties delivered upon selection of the correct odd stimulus window fell into a tray where they remained until the end of the test session. After six consecutive correct trials or 36 trials were completed testing ended. On correct trials the experimenter said 'Yes, that's right' and otherwise: 'No, try again next time'.

Criterion measures. An instrumental criterion of six consecutive correct trials was employed and designated criterion I.

A confrontation trial was presented after the subject had reached criterion I, i.e. either three like stimuli (ear, ear, ear) or three different stimuli (red, yellow, blue) were shown. Some subjects expressed surprise at these displays in which oddity no longer served as a clue to reward and said that they could not solve the problem. Others, without verbalizing the oddity principle, continued to select a window and thus might be deemed without understanding. The latter were asked to explain their choice. Responses indicating understanding of oddity were positively scored, regardless of elicitation procedure. Criterion II was met by a positive verbal response to confrontation.

Results

Criteria I and II. Table 2 presents a summary of the numbers of subjects who achieved criteria I and II as a function of CA. The verbal criterion emerged as the more stringent requirement over all ages and within particular age groups as a majority of subjects first reached criterion I in group 3 (mean CA 65 months), but only in group 4 (mean CA 78 months) did a majority satisfy the verbal criterion II.

Table 2. *Number of subjects reaching criteria I and II as a function of age*

	Groups ($n = 16$) ...	1	2	3	4	5	Total
Day 1	Reaching criterion I	0	3	9	10	12	34
	Reaching criterion II	0	1	4	9	11	25
Day 2	Reaching criterion I	1	4	13	12	15	45
	Reaching criterion II	0	1	6	10	14	31

There is a developmental shift not only in the numbers able to satisfy criterion II, but in the type of explanation offered. The younger subjects of groups 2 and 3 employ 'same' justifications (i.e. 'these two are the same' or 'there are two of these', pointing to the identical stimuli). Indeed, 75 per cent of the verbal responses of these subjects can be classified as 'same' explanations. In group 4 only 42 per cent of the justifications are classified as 'same' explanations, while 37 per cent of the replies stress 'difference' (i.e. 'this one is different' or 'there is only one of these', pointing to

the discrepant stimulus). The proportions reverse in group 5, where there are only 8 per cent 'same' explanations and 'difference' justifications rise to 68 per cent.

The effect of stimulus conditions is seen in Table 3 where criterion I performance alone is presented according to stimulus type. The instrumental criterion is achieved by the younger subjects and met more often by all ages when either of the form stimuli precede colour.

Qualitative analysis. Inspection of the data from which Table 2 is derived indicates that all subjects satisfying both criteria on day 1 also reach both criteria on day 2. The day 2 behaviour of the nine subjects passing criterion I but failing the verbal criterion II is of special interest. There are nine such subjects. Three of the nine, all from groups 3 and 4, pass the verbal criterion II on day 2. None of the remaining six achieves criterion II and two fail to reach criterion I again on day 2. The remaining four subjects do again reach criterion I but three require more trials on day 2 than on day 1. Thus, of the six subjects to meet criterion I but unable to satisfy a verbal criterion, only one subject shows improvement on day 2.

Table 3. *Number of subjects reaching criterion I on each stimulus array*

Groups ...	2		3		4		5		Total	
Day ...	1	2	1	2	1	2	1	2	1	2
Experimental conditions										
Geometric form/Colour	1	1	2	2	3	4	3	4	9	11
Life form/Colour	2	3	4	4	3	4	3	4	12	15
Colour/Geometric form	0	0	1	3	2	2	3	4	6	9
Colour/Life form	0	0	2	4	2	2	3	3	7	9
Total	3	4	9	13	10	12	12	15	—	—

Quantitative analysis. The number of subjects responding to the odd cue on the initial trial on day 1 reveals a further developmental trend. Only 6, 19 and 12 per cent respectively of groups 1, 2 and 3 make their initial choice to the odd stimulus. On the other hand, 56 and 75 per cent of subjects in groups 4 and 5 respond initially to the odd cue.

Parametric analysis based upon error scores yields further information of the effects of CA. An analysis of variance was carried out on the logarithmic transformation $(x+1)$ of total errors per subject on days 1 and 2. Data from group 1 were omitted as preliminary analysis showed chance level performance throughout and no improvement over trials. A $4 \times 4 \times 2 \times 2$ design was employed with four age levels (groups 2, 3, 4 and 5), four experimental conditions (geometric forms/colour, life forms/colour, colour/geometric forms and colour/life forms), two sexes and two administrations of the material (days 1 and 2). The first three variables are between subject main effects and the administration variable which is designated repetition, a within subjects effect.

Four comparisons were significant: age ($F = 7{\cdot}21$; d.f. = 3, 32; $P < 0{\cdot}001$), experimental conditions ($F = 2{\cdot}96$; d.f. = 3, 32; $P < 0{\cdot}05$), repetition ($F = 33{\cdot}82$; d.f. = 1, 32; $P < 0{\cdot}001$) and the repetition × age interaction ($F = 3{\cdot}50$; d.f. = 3, 32; $P < 0{\cdot}05$). Geometric means of error to criterion I are presented in Table 4. Errors decrease significantly with age and with practice from day 1 to day 2, but the inter-

action indicates that day 2 performances improve more with increasing age. An additional analysis of variance which was conducted on the difference scores generated by subtracting each subject's error scores on day 2 from his error score on day 1 lends further support. The only significant comparison was the main effect age, which reached significance at the 5 per cent level. Thus even when their superior day 1 performance is considered, the older subjects are showing superior improvement on day 2.

Table 4. *Geometric mean errors to criterion I as a function of age and experimental conditions*

Groups ...	2		3		4		5	
Days ...	1	2	1	2	1	2	1	2
Experimental conditions								
Geometric form/Colour	21	16	17	10	13	7	9	4
Life form/Colour	18	14	16	11	10	8	14	3
Colour/Geometric form	23	21	19	14	15	14	9	5
Colour/Life form	20	21	17	15	16	12	10	4

Discussion

It is clear that the criteria employed in measuring oddity learning determine the age at which a majority of subjects within a group are reported to have achieved success. Direct comparison can be made with other studies employing an instrumental criterion with normal subjects. The 19 per cent success on criterion I of 4-year-old subjects on day 1 and 25 per cent on day 2 approximates more closely to the single administration, 17 per cent success level reported by Lipsitt & Serunian (1963) than the 42 per cent success ratio described by Gollin & Shirk (1966). Furthermore, subjects in group 5 are not achieving the near errorless performance reported by Gollin & Shirk. Failure to restrict CA and MA overlap between groups, as well as small samples, may have produced inconsistent results in earlier studies. It is suggested that the restriction of CA and IQ range in the present study renders the developmental trend reported here more representative of children of average ability.

The developmental trend from 'same' explanations to 'difference' explanations is especially interesting in so far as subjects in groups 2 and 3 had been successful on the Binet oddity items requiring them to identify the odd stimulus. They comprehend though do not produce the 'difference' label. This shift has also been reported by Saravo & Gollin (1969). The increase in initial oddity responding with age offers further support for a developmental shift from sameness to difference in oddity learning.

The facilitation of learning when the form dimension is relevant evidenced by criterion I attainment and the error analysis is consistent with the literature on children's discrimination learning (Jeffrey, 1968) but extends these findings to include the relational oddity problem. The form/colour comparison in the present study is confounded by the presentation of the colour stimuli, not as a figure in a window but filling the entire aperture. Possibly this method of presentation leads to figure–ground confusion. Additional study of colour as a constant figure on a ground would be necessary before firm conclusions could be drawn regarding the differential

difficulty of form and colour in learning the relevant vehicle dimension of a relational oddity solution.

The authors wish to express their appreciation to the staff and children of the Queen's Park Infant School, Brighton, and The Wallands County Primary School, Lewes, for their generous and willing cooperation.

REFERENCES

ELLIS, N. R. & SLOAN, W. (1959). Oddity learning as a function of mental age. *J. comp. physiol. Psychol.* **52**, 228–230.

GOLLIN, E. S. & SHIRK, E. J. (1966). A developmental study of oddity problem learning in young children. *Child Dev.* **37**, 213–217.

HOUSE, B. J. (1964). Oddity performance in retardates. I. Acquisition and transfer. *Child Dev.* **35**, 635–643.

INHELDER, B. & PIAGET, J. (1964). *The Early Growth of Logic in the Child.* London: Routledge & Kegan Paul.

JEFFREY, W. E. (1968). The orienting reflex and attention in cognitive development. *Psychol. Rev.* **75**, 323–334.

LIPSITT, L. P. & SERIUNIAN, S. A. (1963). Oddity-problem learning in young children. *Child Dev.* **34**, 201–206.

LUNZER, E. A. (1968). *The Regulation of Behaviour.* London: Staples Press.

SARAVO, A. & GOLLIN, E. S. (1969). Oddity learning and learning sets in children. *J. exp. Child Psychol.* **7**, 541–552.

SCOTT, K. G. (1964). A comparison of similarity and oddity. *J. exp. Child Psychol.* **1**, 123–134.

SERPELL, R. (1968). Cultural influence on attentional preference for color over form. *Human Development Research Unit Reports*, no. 8.

SMILEY, S. S. & WEIR, M. W. (1966). Role of dimensional dominance in reversal and non-reversal shift behaviour. *J. exp. Child Psychol.* **4**, 296–307.

SUCHMAN, R. G. & TRABASSO, T. (1966*a*). Color and form preference in young children. *J. exp. Child Psychol.* **3**, 177–187.

SUCHMAN, R. G. & TRABASSO, T. (1966*b*). Stimulus preference and cue-function in young children's concept attainment. *J. exp. Child Psychol.* **3**, 188–198.

(*Manuscript received* 21 *July* 1969; *revised manuscript received* 23 *February* 1970)

Br. J. Psychol. (1971), 62, 1, pp. 27–35
Printed in Great Britain

CONSTANT RATIO RULE FOR CONFUSION MATRICES FROM SHORT-TERM MEMORY EXPERIMENTS

By ROBERT RICH

Department of Psychology, Monash University, Clayton, Victoria, Australia

Clarke (1957) has presented evidence that a constant ratio rule holds for auditory signal identification experiments. That is, the subset of a large confusion matrix may be used to predict a smaller confusion matrix based on data gathered under identical conditions. Evidence is presented that the constant ratio rule also holds for confusion matrices derived from short-term memory experiments. The 10 × 10 confusion matrix presented by Conrad (1964) is 'predicted' from the appropriate submatrix of two 20 × 20 matrices, with moderate success. Besides confirming the applicability of the constant ratio rule to short-term memory data, the experiment also provides indirect support for Conrad's (1965*a*) contention that order errors are a consequence of confusions.

In many experiments stimuli and responses are limited to the same experimenter-specified set. When this is the case, the data may be organized into a confusion matrix: a table listing the frequency (or the corresponding proportion) of each possible response to each stimulus. Thus entries in the main diagonal represent correct responses, and off-diagonal cells cover all the possible errors considered within the confines of the experiment. In short-term memory experiments correct responses are often disregarded (e.g. Conrad, 1964), and the main diagonal is accordingly set to zero.

If a confusion matrix is based on a sufficiently large number of observations, the proportions it contains may be considered as probability estimates. This suggests that if two experiments are identical in every way except in the size of the stimulus/response set used, and the stimulus/response set of one experiment is a proper subset of that of the other, then proportions in the smaller confusion matrix should be the same as for the corresponding submatrix from the larger experiment. That is to say, the presence or absence of other alternatives should not affect the relative frequency with which a particular response is given to a particular stimulus. Clarke (1957) termed this relationship the 'constant ratio rule'. He presented evidence indicating that the constant ratio rule holds for auditory signal identification tasks (Clarke, 1957). Using different groups of unpractised subjects in the two experiments increased the error of prediction from the submatrix of the larger experiment to the matrix of the smaller experiment, but there was no systematic deviation from the rule (Clarke & Anderson, 1957).

It should not be assumed that the same relationship holds for confusion matrices derived from experiments on short-term memory. In a signal identification task, stimuli are presented singly while in memory experiments it is usual to present several stimuli per trial. Thus the distribution of errors is more complexly determined in the latter case. A relationship which holds for signal identification need not be true for memory data. However, it is often useful to be able to compare confusion matrices based on different, but overlapping stimulus/response sets. Thus it is of interest to demonstrate that the constant ratio rule also holds for confusion matrices

derived from short-term memory experiments. However, the large amount of data required tends to make such experiments difficult to carry out. If small stimulus sets are used, within-trial repetitions are necessary but these tend to change the subject's task (Conrad, 1965*b*).

Two large confusion matrices have been prepared from data from an experiment which is as yet incomplete. This is a short-term memory experiment, very similar in procedure to the one reported by Conrad (1964). The stimuli used included the 10 letters used by Conrad, and 10 other consonants. Thus it was possible to establish the validity of the constant ratio rule for short-term memory confusion matrices by comparing the appropriate submatrix of each of these large matrices with Conrad's data. This is the purpose of the research reported here.

METHOD

There were 335 undergraduate subjects from introductory psychology courses at Monash University. They were tested in 13 groups of about 20 each, and two larger groups of about 30 each.

The stimulus/response set consisted of the 20 consonants excluding Y. Six letters were shown on each trial. Stimuli were chosen at random, with the constraints that no letter was repeated on any trial, or for any serial position within a block of 10 trials. Each group was presented with five practice trials, followed by six blocks of 10 trials. There was a short rest after the third block.

The stimulus sequences were drawn in black on an acetate roll, using a Nestler No. 7 lettering stencil. A Beseler 'Porta-Scribe' overhead projector, which was modified by the attachment of a constant speed motor, geared to pull the film at a rate of about 2 in. per sec., was used to present the stimulus sequences. An adjustable mask covered most of the top of the projector, allowing the exposure of a single letter at a time. The stimuli appeared on a white projection screen, moving from right to left as viewed by the subject. Each letter was visible for approximately 1·2 sec.

Each sequence of six letters was preceded by an asterisk. On every trial, the experimenter gave a verbal warning, started the motor, then stopped it as soon as the last letter disappeared. The stopping of the motor was the indication that subjects could start responding.

Responses were in writing, one trial to a page of prepared booklets. Subjects were instructed to commence writing without delay, to attempt to reproduce the six letters in the correct order, and to respond to every letter, guessing if necessary.

Scoring. A response was scored as an error unless it was the letter presented in the corresponding serial position of that trial. That is, scoring was strictly with regard to position in the sequence.

Two 20×20 confusion matrices were constructed. The criterion for inclusion in the first was the same as the one used by Conrad (1964); an error was included if it was the only error on that trial. This eliminates ambiguity concerning the stimulus for which the response was intended (a majority of errors in this type of experiment are order errors). However, this was at the expense of discarding most of the data. Accordingly, a second confusion matrix was constructed by including all errors. Each response was interpreted as being meant for the stimulus in the corresponding serial position. Conrad (1965*a*) argued that order errors are a consequence of the systematic confusions among stimuli, and the strategy imposed on the subject by the task. If this is the case, the second confusion matrix should show the same general trends as the first, although it may be expected to contain more errors due to guessing.

The first matrix will be referred to as the 'single-errors' matrix, the second as the 'all-errors' matrix.

RESULTS

The complete 20×20 confusion matrices are not relevant to the present paper, which is concerned with the correspondence of the appropriate 10×10 submatrix from each with the confusion matrix reported by Conrad (1964). Conrad's matrix is

shown in Table 1. Tables 2 and 3 are the corresponding submatrices from the single-errors and all-errors matrices respectively.

One way of examining the correspondence between two such matrices is by calculating the Spearman rank order correlation coefficient. This was done and a rho of 0·74 was obtained between the single-errors matrix and Conrad's. The corresponding value for the all-errors matrix was 0·76. With $n = 90$, both these values were highly significant (Ferguson, 1966).

Table 1. *Short term memory confusion matrix over a stimulus response set of 10 letters*

(From Conrad (1964), reproduced by permission of the author and publisher.)

	Stimuli									
Responses	B	C	P	T	V	F	M	N	S	X
B	.	18	62	5	83	12	9	3	2	0
C	13	.	27	18	55	15	3	12	35	7
P	102	18	.	24	40	15	8	8	7	7
T	30	46	79	.	38	18	14	14	8	10
V	56	32	30	14	.	21	15	11	11	5
F	6	8	14	5	31	.	12	13	131	16
M	12	6	8	5	20	16	.	146	15	5
N	11	7	5	1	19	28	167	.	24	5
S	7	21	11	2	9	37	4	12	.	16
X	3	7	2	2	11	30	10	11	59	.
Column total	240	163	238	76	306	192	242	230	292	71
Grand total		2050								

Table 2. *Submatrix corresponding to Table 1, from the 'single-errors' matrix*

	Stimuli									
Responses	B	C	P	T	V	F	M	N	S	X
B	.	13	13	2	29	4	6	0	4	4
C	5	.	7	8	12	2	1	2	9	2
P	43	24	.	10	13	14	6	5	3	3
T	14	14	38	.	5	10	2	4	5	13
V	54	11	15	7	.	11	4	3	3	2
F	4	2	4	4	13	.	4	0	23	20
M	3	2	2	2	1	3	.	62	4	4
N	4	7	5	3	7	9	122	.	8	2
S	3	11	0	1	0	40	3	6	.	48
X	2	8	1	2	6	11	1	2	21	.
Column total	132	92	85	39	86	104	149	84	80	98
Grand total		949								

The Spearman rank-order correlation coefficients listed in Table 4 provide a more rigorous test of correspondence. The first column of figures refers to the comparison of the single-errors submatrix with Conrad's matrix. The second column contains correlation coefficients referring to the all-errors submatrix *v.* Conrad's data. Each number expresses the rank correlation for a given stimulus letter, over all possible erroneous responses, i.e. each is based on nine pairs of values. With $n = 9$, a rho of 0·60 is significant at the 0·05 level (Ferguson, 1966). Under an overall null hypothesis

of no association, about one rho in 20 may be expected to have a value of 0·60 or above. In fact, seven of the 10 values in each column are greater than 0·60. Thus the null hypothesis was rejected in both cases.

Table 3. *Submatrix corresponding to Table* 1, *from the 'all-errors' matrix*

	Stimuli									
Responses	B	C	P	T	V	F	M	N	S	X
B	.	31	88	32	120	15	30	10	19	36
C	24	.	77	45	113	33	8	24	49	12
P	172	116	.	62	93	42	45	24	35	17
T	79	63	90	.	43	27	11	22	28	22
V	167	97	67	59	.	336	20	27	45	27
F	15	36	17	14	340	.	33	24	259	119
M	30	14	31	11	31	30	.	153	26	22
N	19	32	36	24	32	45	310	.	37	15
S	28	48	29	10	54	269	24	27	.	192
X	27	18	21	8	27	125	16	13	155	
Column total	561	455	456	265	853	922	497	324	653	462
Grand total		5448								

Table 4. *Spearman rank-order correlation coefficients for each stimulus, comparing corresponding columns of Table* 2 *with Table* 1 (*column* 1), *and of Table* 3 *with Table* 1 (*column* 2)

(A rho of 0·6 is significant at the 0·05 level.)

Stimulus	Rho, single-errors	Rho, all-errors
B	0·853	0·767
C	0·725	0·775
P	0·817	0·750
T	0·771	0·779
V	0·738	0·633
F	0·578	0·745
M	0·354	0·300
N	0·400	0·458
S	0·841	0·850
X	0·638	0·321

In each column, two of the three non-significant correlation coefficients are those for M and N. Observation of Tables 1–3 suggests the reason for this. M and N form a subunit, being highly confusable with each other, but with no other stimulus. Thus it is possible that when a stimulus M is reported as some letter other than M or N, this is through guessing rather than through the utilization of a partially lost stimulus trace. (The same argument may be stated for N as the stimulus.) In such a case, a significant correlation would scarcely be expected. Similarly, for X, the lack of significance of one rho and the marginal value of the other are probably due to the low error rate and even spread across response alternatives for that letter in Conrad's matrix. As evidence for this interpretation it may be mentioned that the rank order correlation for X between Tables 2 and 3 is 0·771.

Clarke (1957) tested the adequacy of the constant ratio rule by plotting the proportion in each cell of the smaller confusion matrix against the proportion in the corresponding cell of the submatrix from the larger experiment. This was done in the

present study, using both the single-errors and all-errors submatrices as predictors (Figs. 1 and 3 respectively). Unfortunately there is no convenient statistic for assessing the degree of fit to the prediction. There seems to be no systematic trend away from the constant ratio rule in either graph, although the points are rather widely scattered, especially for the higher proportions.

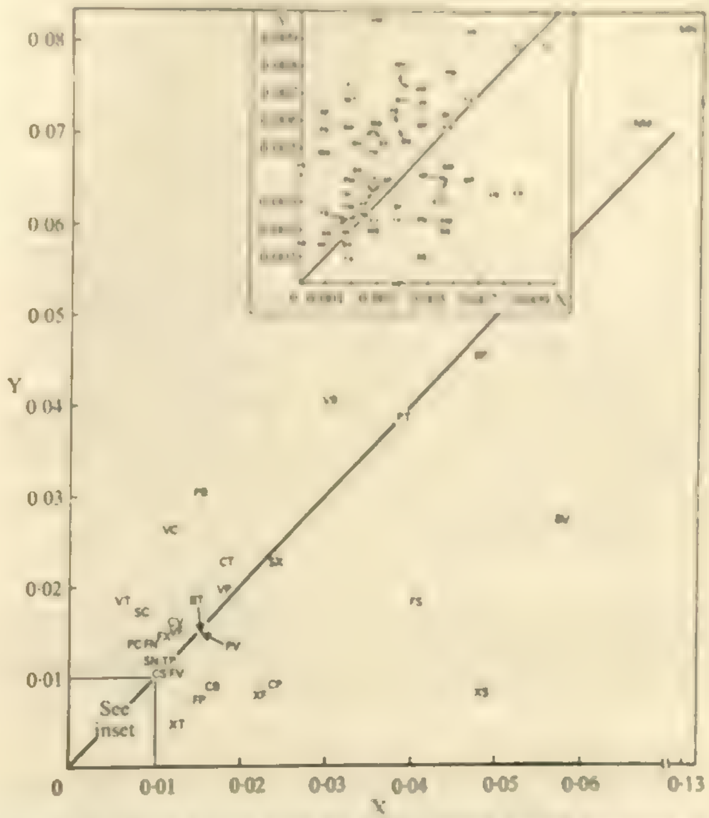


Fig. 1. Proportion of errors in cells of Table 1 (Y), plotted against proportion of errors in corresponding cells of Table 2 (X). Each letter pair indicates the relevant cell (stimulus-response pair). Where points are in close proximity, tip of arrow indicates location.

Some of this scatter is due to a bias in the calculation of the points. Correct letters are excluded from consideration, and each stimulus has a different error rate from table to table and as compared with other letters within a table. In effect, there are unequal stimulus frequencies. Therefore it is appropriate to apply a correction, as suggested by Miller & Nicely (1955). Figs. 2 and 4 are plots of such corrected values, for single-errors and all-errors respectively. Each point on these graphs is calculated by dividing the relevant cell frequency by the column total for the stimulus. Thus each graph may be considered as the composite of 10 independent graphs, each of which is composed of nine points. The 90 points may be considered together, or as 10 separate tests of the constant ratio rule. In either case, the improvement from Figs. 1 and 3 is quite noticeable.

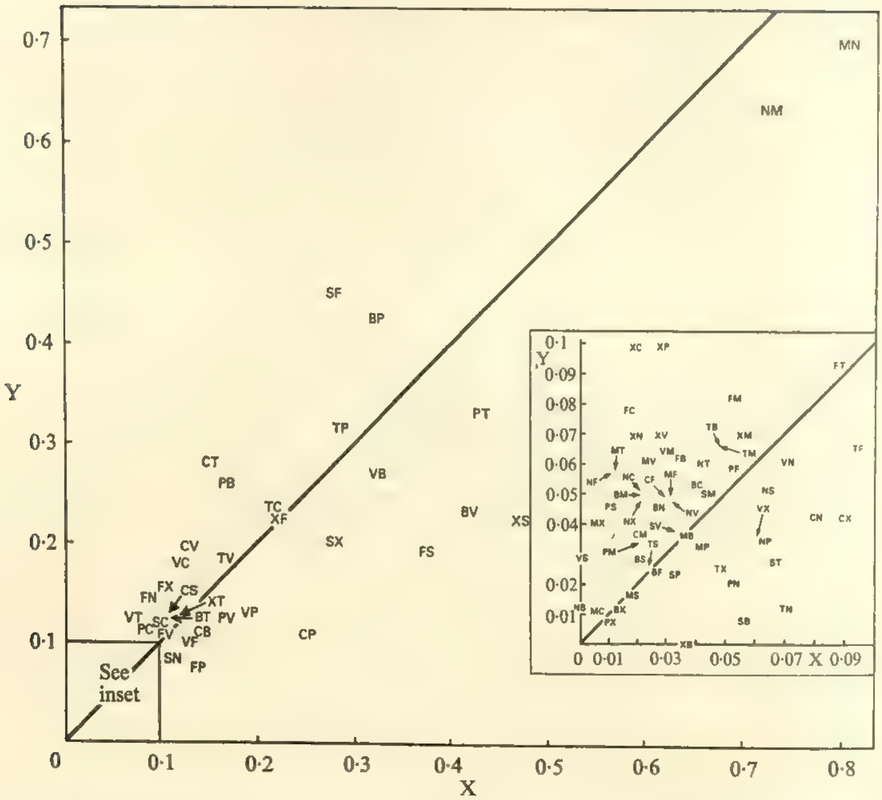


Fig. 2. Proportion of errors, corrected for differing error rates per stimulus from cells of Table 1 (Y), plotted against proportion of errors, corrected for differing error rates per stimulus from corresponding cells of Table 2 (X). Letter pair identifying each point indicates the relevant cell. Where points are in close proximity, tip of arrow indicates location.

Discussion

There are at least two possibly important differences between Conrad's (1964) experiment and the one from which the two larger matrices were derived. First, Conrad used British subjects, while the subjects for the present experiment were almost all Australians. This certainly involves systematic differences in pronunciation, which could be expected to influence the distribution of short-term memory errors. Second, Conrad presented stimuli at the rate of 80 per sec., while in the present experiment the exposure rate was 50 stimuli per sec. This could have led to differences in rehearsal and retrieval strategies used by the subjects in the two experiments. These differences clearly violate one of the assumptions of the constant ratio rule, that the two experiments are identical in all respects except for the stimulus/response set used.

The relative scarcity of data has also weakened whatever relationships may exist between the two experiments. Some of Conrad's frequencies are too low to yield reliable probability estimates. As already mentioned, this may have led to the low values of rho for the stimulus X. The same criticism applies even more strongly to the single-errors submatrix. A total of about 2700 errors were distributed over 380

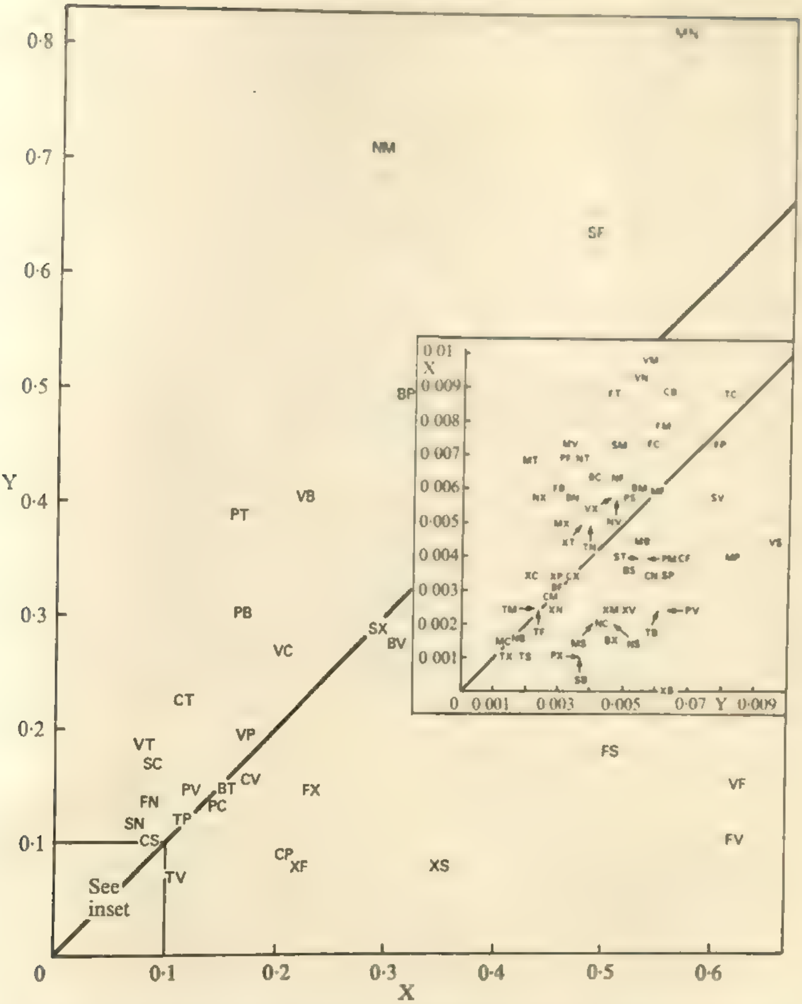


Fig. 3. Proportion of errors in cells of Table 1 (Y), plotted against proportion of errors in corresponding cells of Table 3 (X). Each letter pair indicates the relevant cell. Where points are in close proximity, tip of arrow indicates location.

cells—an average of only 7·1 per cell. Even though the selected submatrix contained many of the high-frequency confusions, the average for the 90 selected cells was only about 10·5 entries per cell. Thus the ordering of the low-frequency cells is unlikely to reflect any systematic tendency. Had more data been available for the single-errors matrix and for Conrad's matrix, the correspondence to the constant ratio rule would have been better.

The all-errors matrix does not suffer from this shortcoming. However, its method of scoring introduces another difference, further violating the assumption underlying the constant ratio rule. For this reason, it is doubtful whether further data would improve results based on this matrix.

When the above considerations are kept in mind, it is surprising that the constant ratio rule should hold as well as the results indicate. Figs. 2 and 4 show unmistakable

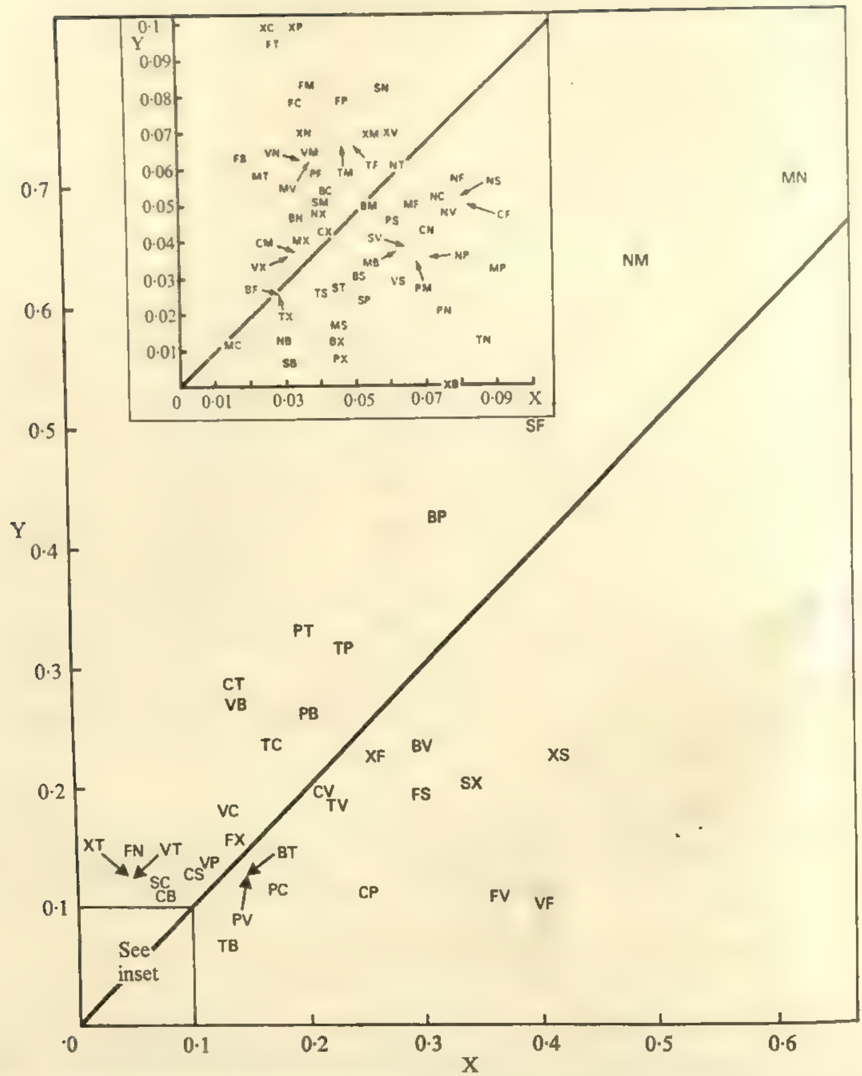


Fig. 4. Proportion of errors, corrected for differing error rates per stimulus from cells of Table 1 (Y), plotted against proportion of errors, corrected for differing error rates per stimulus from corresponding cells of Table 3 (X). Letter pair identifying each point indicates the relevant cell. Where points are in close proximity, tip of arrow indicates location.

trends toward the 45° line which defines the relationship postulated by the constant ratio rule. It is impossible to make quantitative comparisons with Clarke & Anderson's (1957) graphs, but it is not implausible that if the above sources of error were eliminated, the 'predicted' proportions would correspond as well to those of the smaller matrix as in Clarke & Anderson's paper.

Unfortunately, Clarke & Anderson did not use rank correlational methods for assessing their findings. Had they done so, a quantitative comparison might have been possible. As it is, one has to be content with the conclusion that the constant ratio rule holds to some extent for confusion matrices derived from short-term

memory experiments, and, possibly, it holds as well as for signal identification experiments (provided that the assumptions of the rule are not violated and the matrices contain sufficient data).

As an incidental finding, the relative success of the all-errors matrix in predicting a matrix obtained by a 'single-errors' scoring criterion provides some support for Conrad's explanation of the occurrence of order errors in short-term memory experiments (Conrad, 1965*a*). According to Conrad, order errors tend to occur because, after a few trials, the subject knows that repetitions are not likely within adjacent serial positions of a trial. If two highly confusable letters are adjacent or nearby, there is a high probability that an erroneous response to one will be the other. Thus in order to avoid repetition, the subject has to choose a response other than the correct letter for the second serial position in question. As confusability tends to be mutual, this is likely to be either the stimulus from the previous serial position, or a stimulus which is confusable with both. This gives rise to transpositions and semi-transpositions respectively.

Over 50 per cent of errors in short-term memory experiments are order errors (Conrad, 1965*a*; Rich, 1967). Conrad's (1964) scoring method excludes such order errors, while they are included in the all-errors matrix. The correspondence between the two is hard to account for unless there is a close relationship between order errors and confusions. Of course, the causal sequence could be the opposite to that described by Conrad (1965*a*), but then singly occurring errors become hard to explain. Or both kinds of errors could be due to some other factor.

References

CLARKE, F. R. (1957). Constant-ratio rule for confusion matrices in speech communication. *J. acoust. Soc. Am.* **29**, 715–720.

CLARKE, F. R. & ANDERSON, C. D. (1957). Further test of the constant-ratio rule for speech communication. *J. acoust. Soc. Am.* **29**, 1318–1320.

CONRAD, R. (1964). Acoustic confusions in immediate memory. *Br. J. Psychol.* **55**, 75–84.

CONRAD, R. (1965*a*). Order error in immediate recall of sequences. *J. verb. Learn. verb. Behav.* **4**, 161–165.

CONRAD, R. (1965*b*). The role of the nature of the material in verbal learning. *Acta psychol.* **24**, 244–252.

FERGUSON, G. A. (1966). *Statistical Analysis in Psychology and Education.* New York: McGraw-Hill.

MILLER, G. A. & NICELY, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. acoust. Soc. Am.* **27**, 338–352.

RICH, R. (1967). Unpublished M.Sc. thesis, University of Sydney, Australia.

(*Manuscript received* 26 *January* 1970)

Br. J. Psychol. (1971), **62**, 1, pp. 37–40
Printed in Great Britain

STIMULUS SIMILARITY AND TRANSFER IN LONG-TERM PAIRED-ASSOCIATE LEARNING

BY ALEX McGLAUGHLIN AND H C A DALE
Department of Psychology, University of Hull

Acoustic, semantic, visual and formal similarity were studied in the A-B, A′-B transfer paradigm. Substantial, highly significant, positive transfer was obtained with stimuli that were acoustically or semantically related. For those visually or formally related, the transfer was not significant. It was concluded that the failure to find significant effects of acoustic similarity in previous long-term memory paired associate learning studies has been due to a combination of low levels of acoustic similarity with the use of the relatively insensitive retroactive interference technique.

The Transfer and Retroaction Surface proposed by Osgood (1949) provides a model accounting for all serial and paired-associate learning transfer phenomena. In this the amount and direction of transfer are predicted according to the relationship between the indices of inter-list stimulus and response similarity, but no attempt is made to specify the dimension along which such transfer occurs.

Studies of retroactive interference (RI) with acoustic and semantic similarity between original learning (OL) and interpolated learning (IL) have led to the recent emergence of a simple generalization. When the classical A-B, A′-C interference paradigm is used in long-term memory (LTM), semantic similarity increases RI (e.g. Baddeley & Dale, 1966), whereas acoustic similarity does not (e.g. Dale & Baddeley, 1969). These studies have employed adult subjects. Hence it would appear that words are coded semantically in LTM by adults.

Some evidence from the field of education runs counter to this conclusion. Thus many spelling errors can be attributed to acoustic confusions (e.g. Burt, 1921). Phonic training is important and useful as a didactic device when children are being taught to read (Schonell, 1945). Finally, the success of the initial teaching alphabet (i.t.a.) can be attributed to the consistency with which it signals sounds. Thus there is a strong indication that normal children can and do use the acoustic properties of words when learning to read and write.

A number of explanations for this apparent contradiction can be advanced. It could result from a developmental factor, with reliance on acoustic coding replaced by reliance on semantic coding as age increases. Alternatively, the absence of acoustic similarity effects in the studies quoted above may be due to artifacts of the laboratory situation.

A comparison between the PAL retroactive or negative transfer paradigm of the psychological laboratory and the positive transfer techniques used in classroom instruction suggests that the former could be relatively insensitive. Subjects in the laboratory may be able to successfully minimize RI by suppressing generalization along some dimensions, e.g. acoustic, and thus produce the false impression that generalization was entirely absent. In a positive transfer situation where it is to the subject's advantage to use generalization he is more likely to reveal its presence.

The present study examines the possibility that when a positive transfer paradigm

is used in PAL with adult subjects, stimulus similarity on both acoustic and semantic dimensions will enhance transfer. The transfer possibilities of formal and visual similarity are also examined, although formal similarity has been found to be ineffective elsewhere (e.g. Baddeley, 1966*b*).

METHOD

Design. A transfer design of the classic A-B, A′-B form was used for eight independent groups. All groups received the same original learning (OL) list of 16 pairs and then one of eight transfer learning (TL) lists. Responses comprised the series of two-digit numbers 13–29 with the exclusion of 22. The OL stimuli were 16 words chosen from the Kent–Rosanoff (1910) norms having a Thorndike–Lorge (1944) G frequency exceeding one per four million. (Their mean value was 71·5.) Eight intermediary lists of 16 words were then prepared, each bearing a single similarity relationship to the list of OL stimuli, and from these the eight 16-pair transfer lists were derived. The first intermediary list contained 16 words chosen so that each was formally similar to one of the OL words. Care was taken to ensure that it was in no way similar to any other OL word, a principle which was also applied to the other lists. The level of formal similarity employed was constant. Each transfer word was constructed from the same letters as the corresponding OL word with the exception of one letter. Thus *fruit* corresponded to *refit.* The semantically similar intermediary list contained words listed in *Roget's Thesaurus* as synonyms of those on the OL list. The visually similar mediating list contained words whose outline shape, when printed in lower-case lettering, was identical to those on the OL list. In preparing this material, the letters b, d, h, k, l and t were treated as identical, as were f, g, j, p, q, y (our initial matching was for handwritten script in which the f came below the line and we modified that stamp accordingly), and a, c, e, i, m, n, o, r, s, u, v, w, x, z. The acoustically similar mediating list contained words identical to those on the OL list on all but one phoneme. The last four intermediary lists served as a source of control words. These were dissimilar to all those words on the OL list on all the four dimensions of similarity being considered.

The mean frequency of occurrence in the *Thorndike–Lorge* G count of words in each list was between 31 and 34 per million.

Each transfer list was then constructed by selecting randomly, without replacement, two words from each of the eight intermediary lists with the constraint that every OL stimulus was paired with one word in each transfer list. Thus every word on the OL list was represented by four other related words; one formally, one semantically, one visually, one acoustically and by four corresponding words from the control intermediary lists. The OL stimulus *heavy*, for example, was represented in this way by the words *halve*, *weighty*, *dairy*, *heady*, *probe*, *fast*, *under* and *lest*, respectively. Each subsequently appeared in just one of the transfer lists. At this stage all materials were exhaustively checked to ensure that the only similarity relationships that existed between an OL stimulus and a transfer list stimulus were those specified, all other intra-list and inter-list similarities being eliminated.

Procedure. Eight separate groups of subjects learned the OL list for eight trials and then one of the eight transfer lists for one trial only.

The subjects were tested in groups varying in size from 17 to 32 members. The stimulus material was printed in black 2 in. high lower-case letters on white card 5 × 21 in. in size, using rubber stamps. The cards were held up by the experimenter so as to be clearly visible to all subjects. The responses were written in specially printed grids with one line for each trial. The limits of the response-set (namely, 13–29) were displayed on a blackboard throughout the session. On every trial the stimulus-response pairs were displayed successively for 4 sec. each. In the subsequent test, each stimulus was presented alone for 4 sec. The inter-item interval was 2 sec. during presentation and 1 sec. during the test. There was an interval of 10 sec. between presentation and test within each trial, and an interval of 10 sec. between trials. A metronome was used to ensure correct pacing. Immediately following the eighth test trial all subjects were required to transcribe a message consisting of random digits for 3 min. This had been pre-recorded at a rate of 1 digit per sec. and served as a STM control (Baddeley, 1966*a*). The subjects were young enlisted men. In their instructions, the stimulus-response pairs were introduced as challenges and passwords. Following OL and the digit transcription task, they were told: 'The

enemy has discovered the passwords. These must therefore be changed and it is necessary to learn a new set. In this particular instance the same set of numbers will again serve as passwords but the challenges will be new words.' The task, as before, was to learn to respond rapidly with the correct password to each challenge.

Results

Analysis was confined to those subjects who reached the criterion of 16 correct responses on the last OL trial. The number reaching this criterion in each group varied from 10 to 17. It is necessary to equate group size for analysis, to ensure that equal weight is given to all stimulus pairs irrespective of the level of their similarity. Hence 10 subjects were randomly selected to represent each group.

Table 1. *Mean level of TL performance*

Stimulus relationship	Transfer learning: Raw score (max = 160)	Transfer learning: Per cent
Semantic	116	72·5
Acoustic	110	68·75
Formal	87	54·375
Visual	69	43·125
None (C_1)	80	50·0
None (C_2)	74	46·25
None (C_3)	74	46·25
None (C_4)	67	41·875

On the transfer list all subjects had the opportunity to learn two pairs under each treatment. With eight groups, each of 10 subjects, 160 correct responses would be scored for perfect TL performance. The raw scores obtained are shown in Table 1. Also shown is the mean level of TL performance expressed as a percentage for each of the four similarity treatments and the four control, or non-similarity, treatments.

A two-factor, 8×8 (groups $\times$ treatments) analysis of variance, using raw scores on the transfer list as data, revealed a highly significant effect of treatments ($F = 5{\cdot}11$; d.f. = 7, 56; $P < 0{\cdot}001$), but no significant effect of groups ($F = 1{\cdot}94$; d.f. = 7, 8; $P > 0{\cdot}05$) nor of the groups $\times$ treatments interaction ($F = 0{\cdot}85$; d.f. = 49, 56; $P > 0{\cdot}05$).

Scheffé's test for multiple comparisons (Edwards, 1950) had been planned between the group of four control means and each of the similarity treatment means. This showed that performance on the semantically similar material was significantly better than that on the control material beyond the 1 per cent level ($MS_D = 115{\cdot}7$; $K = 8$; $V = 105$), while performance on the acoustically similar material was better at just less than the 2·5 per cent level ($MS_D = 82{\cdot}1$; $K = 8$, $V = 105$; $P < 0{\cdot}05$). Neither the formally similar nor the visually similar materials significantly, affected performance ($MS_D = 10{\cdot}97$ and $1{\cdot}41$ respectively; $K = 8$, $V = 105$).

Discussion

In this experiment, positive transfer was shown to result from acoustic similarity between OL and TL stimuli. Such a clear effect of acoustic similarity in LTM has not previously been demonstrated. Semantic similarity, which influences RI

(Baddeley & Dale, 1966), also influenced transfer strongly, whereas visual and formal similarity did not.

Sassenrath & Yonge (1967) found positive transfer for homonyms (by which they mean homophones, e.g. *weak* and *week*), in the A-B, A′-B paradigm, although not at a statistically significant level. The degree of acoustic similarity in the study reported here did not reach the level of identity, but a concurrent study, in preparation, concerning itself solely with acoustic identity of stimuli in the A-B, A′-C, RI paradigm, demonstrates highly significant effects. The failure of previous studies (e.g. Dale & McGlaughlin, 1968) to demonstrate the effects of acoustic similarity for PAL in LTM is thus attributable to a combination of two factors: the comparatively low level of acoustic similarity of the word pairs generally chosen for study, and the use of the relatively insensitive RI technique.

The authors gratefully acknowledge the financial support of the Medical Research Council. The work undertaken by the senior author is to be offered in part fulfilment of the requirements of the Ph.D. degree in the University of Hull.

REFERENCES

BADDELEY, A. D. (1966*a*). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Q. J. exp. Psychol.* **4**, 302–309.

BADDELEY, A. D. (1966*b*). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Q. J. exp. Psychol.* **4**, 362–365.

BADDELEY, A. D. & DALE, H. C. A. (1966). The effect of semantic similarity on retroactive interference in long- and short-term memory. *J. verb. Learn. verb. Behav.* **5**, 417–420.

BURT, C. (1921). *Mental and Scholastic Tests.* London: Staples Press.

DALE, H. C. A. & BADDELEY, A. D. (1969). Acoustic similarity in long-term paired-associate learning. *Psychonom. Sci.* **16**, 209–211.

DALE, H. C. A. & MCGLAUGHLIN, A. (1968). Acoustic similarity and retroactive inhibition in paired-associate learning. *Psychon. Sci.* **13**, 225–226.

EDWARDS, A. L. (1950). *Experimental Design in Psychological Research.* New York: Holt, Rinehart & Winston.

KENT, G. H. & ROSANOFF, A. J. (1910). A study of association in insanity. *Am. J. Insan.* **67**, 317–320.

OSGOOD, C. E. (1949). The similarity paradox in human learning: a resolution. *Psychol. Rev.* **56**, 132–143.

SASSENRATH, J. M. & YONGE, G. D. (1967). Meaning and transfer of verbal learning. *J. educ. Psychol.* **58**, 365–372.

SCHONELL, F. J. (1945). *The Psychology and Teaching of Reading.* Edinburgh: Oliver & Boyd.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teachers Wordbook of* 30,000 *Words.* Teachers College Columbia University.

(*Manuscript received* 21 *February* 1970)

Br. J. Psychol. (1971), 62, 1, pp. 41–51
Printed in Great Britain

REMINISCENCE IN PAIRED-ASSOCIATE LEARNING

By MUKUL K. DEY*

Mount Allison University, Sackville, New Brunswick, Canada

Using lists of six tone word pairs, two degrees of intra-list stimulus similarity were factorially combined with three rest pauses (0, 2 and 4 min.) interpolated after the attainment of criterion 4/6. The recalls on the first two post-rest trials reflect a depressive rather than facilitative influence of rest pause. However, the analyses of post-rest trials to criterion 6/6 have revealed that significant facilitation occurred only when the length of rest pause was 2 min. and stimulus similarity was high. The latter fact proves that a reminiscence effect resembling the Ward–Hovland phenomenon is obtainable in paired-associate learning provided the similarity among stimuli is high enough. The results appear to be best explained by a theory of paired-associate learning which postulates a temporally labile competition-generated inhibitory potential.

In an experiment which is now a classic in human learning and retention, Ward (1937) demonstrated that an experimental group which receives a brief rest interval ranging from ½ min. to 5 min. after partially learning a serial list of nonsense syllables exhibits superior performance following the interpolated rest in comparison with a control group which continues practising without interruption. The phenomenon known as reminiscence was subsequently confirmed in several studies by Hovland (1938*a*, *b*, 1939). A short-time reminiscence effect similar to the one obtained in the serial learning of nonsense syllables can be expected in paired-associate (PA) learning, because of close resemblance between the two rote-learning methods, but the evidence thus far derived does not bear out this expectation. Hovland (1939) failed to notice any significant improvement after interpolating a 2 min. rest pause (RP) in the learning of paired nonsense syllables, though in the same experiment reminiscence was amply produced in the learning of serial lists under comparable conditions. More recently, Riley (1953, 1954, 1957) has reported a series of six experiments, out of which only two (Expts. 2 and 4) have adopted the standard design. In these experiments the interpolation of a 2-min. RP after practice through either six or 10 trials did not have any significant influence upon post-rest performance. The other experiments have actually employed such designs as are normally used in retroaction and proaction studies, and therefore are more suitable for answering questions relating to decrement in retroactive or proactive interference resulting from a brief interval following interpolated learning than for studying reminiscence *per se*. At best these experiments may offer information on the influence of a short RP upon the learning of a PA list which is preceded by another PA list. Strictly speaking, a genuine reminiscence effect is the product of one-list learning and not of interaction between lists. An earlier claim by McGeoch *et al.* (1937) is also to be discounted for two reasons: (1) The sequence of pairs was constant rather than randomly altered between trials, a procedure not normally employed in PA learning. (2) Instead of employing the usually recommended control-group method, this study compared the recall scores by the same group on the criterion trial and on the trial immediately after rest, which method has been experimentally proven to be unsound (Gray, 1940).

* Present address: Department of Psychology, Jersey City State College, Jersey City, N.J., U.S.A.

The attempts to demonstrate facilitation of PA learning by distribution of practice (DP) have failed as often as they have succeeded. Since distributed practice efficacy is in a theoretical sense the sum total of a series of successive reminiscence effects (McGeoch, 1942), the reason for failing to find reminiscence may be the same as, or at least closely related to, the reason behind frequent failures to show DP facilitation in PA learning. A theory presented by Underwood (1961) identifies initial degree of integration among the elements of response terms as a crucial factor determining whether or not DP facilitation may appear. In view of accumulating negative evidence (Jung, 1966; Marshall & Runquist, 1962; Underwood *et al.*, 1964), while there is positive evidence from only one study (Underwood & Schulz, 1961), Underwood & Ekstrand (1967) have lately expressed reluctance to accept this theory. In an alternative explanation, Dey (1966*a*) has emphasized the importance of intra-list stimulus similarity (ISS) as a primary variable in the DP facilitation of PA learning. Clear demonstrations of faster PA learning with increased inter-unit interval (IUI), i.e. the time elapsing between the offset of a response term and the onset of the next stimulus term, by Dey (1966*a*) and with increased inter-trial interval (ITI) by Underwood & Ekstrand (1967) when ISS is high but not when it is low, have strengthened this point of view. It is indeed worthwhile to inquire whether sufficient ISS is likewise an essential condition for reminiscence to appear in PA learning.

METHOD

General procedure

Two levels of ISS—low (LS) and high (HS)—and three rest pauses (0, 2 and 4 min.) gave a 2×3 factorial design to the experiment. The groups treated to the six conditions in this design are named LS-0, HS-0, LS-2, HS-2, LS-4 and HS-4. The control groups (LS-0 and HS-0) practised continuously up to the 6/6 criterion, which means first correct anticipation of all six response terms on a trial. The experimental groups (LS-2, HS-2, LS-4 and HS-4) practised up to the 4/6 criterion, or to a higher criterion if the former happened to be skipped, received the assigned RP, and finally practised again until their performance reached the 6/6 criterion. During the interpolated RPs, the experimenter read out a list of simple words (e.g. good, up, light, day) with readily available opposites (e.g. bad, down, dark, night) at intervals of approximately 8 sec., while the subject responded to these by calling out the opposites as promptly as possible.

Learning material and equipment

The stimuli of the six-unit PA lists were pure tones selected from within the range where the Weber fraction is generally considered to be stable (cf. Woodworth & Schlosberg, 1954). The cycle frequencies of the tones were 410, 575, 805, 1125, 1580 and 2210 for the LS lists (Δ/S ratio $= 0{\cdot}4$) and 410, 490, 590, 710, 850 and 1020 for the HS lists (Δ/S ratio $= 0{\cdot}2$). The same set of six three-letter words were used as the response terms in all lists: *sky*, *car*, *fun*, *dog*, *net* and *lip*.

While the learning material was presented entirely by a tape-recorder, the following additional equipment was employed in tape-recording: a Hewlett–Packard audio-oscillator, a Stoelting voice relay, a Scientific Prototype L.D.R. audio-switch, and three Hunter interval timers. The event and interval durations used in recording were: (1) 0·75 sec. as the tone duration, (2) 3·25 sec. as the anticipation interval, (3) 1·5 sec. as the time taken for each word presentation, and (4) 2·5 sec. as the IUI. It may be noted that all except the tone duration are approximate values. For each ISS level, while using one pattern of S-R pairing, three hour-long tapes were prepared with three different series of randomly derived pair sequences. Each of these tapes was used with four subjects. Where RP was interpolated, list presentation was always resumed from the point where the tape was stopped prior to RP. Note that except for the RPs, when interpolated, the presentation of learning material was continuous, so that ITI was the same as IUI.

Instruction

The following typewritten instruction was read out to the subjects at the outset: 'The material for learning will be presented by means of this tape recorder. As we start, first a tone and afterwards a word will be heard. A brief interval will then pass, after which there will appear a second tone followed by a second word, and so on. The task has six such tone–word pairs. The six pairs will be presented over and over again until your performance reaches a certain level. Note that the tone and the word in each pair will always belong together, but in the course of repetition, the tone–word pairs will always appear in different orders. We want you to form associations between the members of each pair so that at any tone presentation you may be able to recall the word that is paired with it. After the six pairs have been presented once, as soon as a tone is heard, we should like you to call out the word that is to follow. Let me emphasize that you have to be quick in responding, as the time between the presentations of the tone and the word is very short. During the interval elapsing between one pair and the next, please do not engage in any recitation, silent or loud, of what you may have learned but attentively wait for the next tone. Be fully alert and effortful throughout the experiment. Kindly do not discuss the tones and words used in this experiment with your associates in the college as some of them may be our future subjects. If you have any question, let me hear it right now.'

To subjects receiving RPs, the following instruction was read out just before practice was to be resumed: 'We shall now have a second session of the same experiment, which involves the learning of associations between tones and words. The tone–word pairs and also the procedure of responding will be the same as before. However, this time I want you to start responding as soon as task presentation is resumed.'

Subjects

Forty-two male and 30 female undergraduate students with a mean age of 21·8 yr., who were randomly obtained from certain sophomore and junior classes in psychology, the only restriction being that none had taken any course in musical performance or appreciation, were equally divided among the six groups. A 7:5 male–female representation was strictly observed for all groups. Several days before the test on PA learning, a quick test for hearing was administered to each subject. Three subjects were rejected on the basis of this hearing test. Five more subjects were rejected as, in the middle of the experiment, they either reported inadequate motivation or appeared to have misunderstood the instruction. They were replaced by additional subjects drawn from the same classes.

Results

Pre-rest learning

The mean trials required to learn the LS and HS lists to the pre-rest criterion (4/6) are, respectively, 8·36 and 11·08. The increase in effort by 32·53 per cent as a result of a decrease from 0·4 to 0·2 in the Δ/S ratio in the stimulus series is found to be significant at the 0·02 level by a two-tailed test (t = 2·41; d.f. = 70). The mean trials to criterion 4/6 for the group treated to LS are 8·08, 9·25 and 7·75, and those for the groups treated to HS are 11·25, 10·25 and 11·75. Separate one-way classification analyses of variance for these two groups of means have revealed non-significant differences ($F < 1$, d.f. = 2, 33, in both cases). This testifies to equivalence in learning ability among the three groups treated to either condition of ISS.

Post-rest recall

The recall scores on each post-rest trial from 1 to 5 were subjected to a 2×3 covariance analysis (Winer, 1962), using trials to criterion 4/6 as the covariate. The covariance-adjusted as well as unadjusted group means are shown in Table 1. Some of the means for trials 4 and 5 are slightly overestimated, as certain subjects had reached the final criterion (6/6) in three or fewer trials and so their scores on later trials were

counted as 6, though it could very well be that if their performance was carried on further, the later scores in some of the cases would be less than 6 on account of fluctuation of recall. The mean recalls appear to decrease with increase of RP on post-rest trials 1 and 2 with both LS and HS lists. The covariance analyses attest the reliability of these trends by revealing significance in the variances caused by RP ($F = 3{\cdot}16$, d.f. = 2, 65, $P < 0{\cdot}05$; $F = 4{\cdot}07$, d.f. = 2, 65, $P < 0{\cdot}01$), while the lack of significance in the RP × ISS interaction ($F < 1$, d.f. = 2, 65, in both cases) indicates that these trends are unaffected by ISS. Similar covariance analyses of the recall scores obtained with both LS and HS lists on post-rest trials 3, 4 and 5 have revealed significance in neither the effect of RP nor the RP × ISS interaction.

Table 1. *Unadjusted and covariance-adjusted means of correct recalls on five post-rest trials*

	Unadjusted					Adjusted				
Group	1	2	3	4	5	1	2	3	4	5
HS-4	2·42	2·25	3·75	3·92	4·00	2·27	2·24	3·68	3·89	3·93
HS-2	2·75	3·25	3·58	4·00	4·42	2·71	3·25	3·56	3·94	4·40
HS-0	3·25	3·42	3·67	3·67	3·83	3·13	3·41	3·62	3·64	3·78
LS-4	2·17	2·75	3·58	3·67	4·25	2·32	2·76	3·65	3·70	4·32
LS-2	2·83	3·00	3·67	3·83	4·08	2·87	3·00	3·69	3·84	4·10
LS-0	3·33	3·50	3·83	3·75	4·17	3·45	3·51	3·88	3·78	4·22

Post-rest learning

Though the recall data do not furnish any reliable evidence of reminiscence, the relearning data do. The group means of post-rest trials to criterion 5/6 and to criterion 6/6 were covariance-adjusted, using pre-rest trials to criterion 4/6 as the covariate. The unadjusted and adjusted means are shown in Table 2. The curves plotted with the adjusted means (Fig. 1) suggest that reminiscence is (*a*) manifest if post-rest performance is continued long enough, especially up to criterion 6/6, (*b*) considerably greater with the HS than with the LS stimuli, and (*c*) more effectively produced by the 2-min. than by the 4-min. RP.

A 2 × 3 covariance analysis of the post-rest trials to criterion 5/6 has revealed no significant effect whatsoever. However, a similar analysis of the post-rest trials to criterion 6/6 has yielded an F ratio of 4·75 for RP which, with 2 and 65 d.f., is significant between the 0·05 and 0·02 levels. Though the same analysis has produced a non-significant F ratio (1·05) for RP × ISS interaction, the one-way classification covariance analyses carried out separately for the two levels of ISS have revealed a significant difference among the HS-0, HS-2 and HS-4 means ($F = 6{\cdot}79$; d.f. = 2, 32; $P < 0{\cdot}01$), but not among the LS-0, LS-2 and LS-4 means ($F < 1$; d.f. = 2, 32). Further, paired comparisons by covariance analyses were made between HS-0 and HS-2 and between HS-0 and HS-4. A reliable difference is shown in the former comparison ($F = 23{\cdot}99$; d.f. = 1, 21; $P < 0{\cdot}01$), but not in the latter ($F = 1{\cdot}82$; d.f. = 1, 21; $P < 0{\cdot}1$). These statistical findings, while generally upholding the observations based upon an inspection of the curves, unequivocally prove that a reminiscence effect resembling the Ward–Hovland phenomenon can, indeed, be demonstrated in PA learning, using a short rest pause like 2 min., provided the stimuli have sufficiently high similarity.

Table 2. *Unadjusted and covariance adjusted means of post rest trials to the 5/6 and 6/6 criteria*

Group	Unadjusted 5/6	Unadjusted 6/6	Adjusted 5/6	Adjusted 6/6
HS-4	7·83	15·58	7·89	16·35
HS-2	5·50	9·75	5·51	9·95
HS-0	8·50	21·17	8·54	21·74
LS-4	6·67	14·67	6·62	13·92
LS-2	6·42	11·50	6·41	11·32
LS-0	6·92	16·33	6·87	15·71

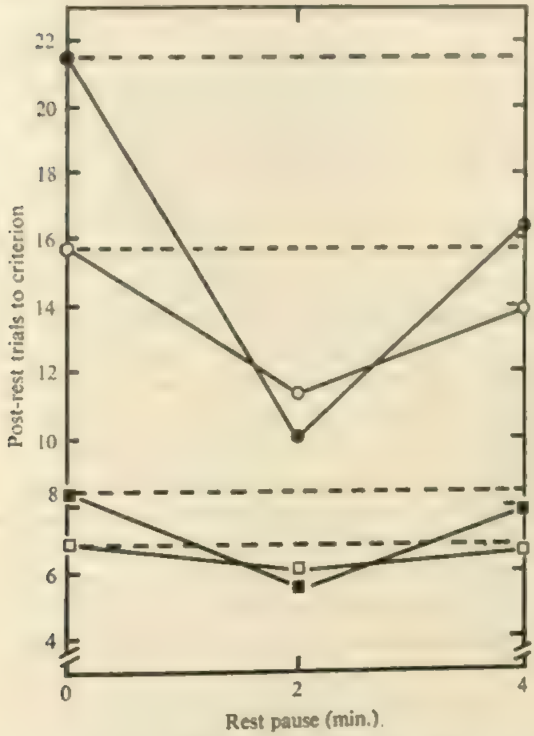


Fig. 1. The curves drawn with solid lines represent the adjusted mean recalls as functions of rest pause and intra-list stimulus similarity, while the horizontal dashed lines represent the control levels. 6/6: ●, HS; ○, LS. 5/6: ■, HS; □, LS.

Discussion

Our finding that recall immediately following RP manifests depression which is in direct proportion to the length of RP is not atypical in the studies of reminiscence in verbal learning. A very similar phenomenon was reported by Archer (1953) and by Melton & Stone (1942). The observations of Melton & Stone (1942) further resemble ours in that the magnitude of the phenomenon was considerably reduced in the next few trials. Depression of recall as the immediate outcome of RP and the positive relationship of the amount of depression to length of RP in the present experiment are

attributable to two factors: (1) The rest-interval activity consisted in verbal associations which, in spite of ease of elicitation, might have inhibited the PA task's responses through unlearning or competition at recall or both. Evidence from serial learning indicating poorer recall after a verbal rest-interval activity than after a nonverbal one (Shipley, 1939) adds credence to the above hypothesis. (2) Irion (1949) has demonstrated with convincing evidence that loss of retention of serial associations following short intervals is partly due to disruption of set, which, being simply the outcome of discontinuity in performance, bears a directly proportionate relationship with the period through which performance is discontinuous. One of the reasons why extreme caution is recommended in the choice of rest-interval activity in any study of short-term reminiscence is to ensure that continuity of the work-set is maintained in some way. The rest-interval activity employed in our experiment no doubt keeps up the rhythm of stimulation and responding, but by directing the verbal associations along a certain specific channel, it perhaps produces obstruction to the continuation of the work-set. However, presumably the process generating reminiscence was intensive enough to overcome the inhibitory influence of the above two factors so that reliable reminiscence subsequently appeared in the form of reduced effort in extended learning.

The optimal conditions for reminiscence in serial learning are massed practice, prerest learning to a moderate criterion, and a brief rest interval like 2 or 3 min. To the writer's knowledge, the literature reports 10 studies that have observed these optimal conditions and at the same time used the number of post-rest trials to a criterion higher than the pre-rest criterion for additional measurement of reminiscence. Of these, only two have demonstrated significant reminiscence in recall but not in extended learning (Buxton & Bakan, 1949; Buxton & Ross, 1949). The remaining studies have either shown reminiscence through both recall and extended learning (Ward, 1937; Hovland, 1938*a*, *b*, 1939)—in fact, all the three studies of Hovland have produced greater reminiscence in extended learning than in recall—or furnished evidence of reminiscence in extended learning while the recall data are inconclusive (Hovland & Kurtz, 1951; Melton & Stone, 1942; Noble, 1950; Shipley, 1939). These suggest the conclusion that the number of post-rest trials required to reach a higher criterion provides a more accurate measurement of reminiscence than the recalls immediately after RP do, and raise the question whether some of the studies of serial learning which failed to obtain reminiscence under the optimal conditions (e.g. Underwood, 1957) might demonstrate it if performance was continued well beyond the first trial after rest.

The three processes which can be readily ruled out as probable causes of reminiscence in the present experiment are fatigue, motivation and rehearsal. The untenability of both fatigue and motivation is well-proven by Hovland's (1938*a*, *b*, 1939) demonstration of significant reminiscence after single-trial presentation of serial lists of nonsense syllables. Whithey *et al.* (1949) report greater post-rest recall when subjects were instructed to rehearse than when they were not. An exactly opposite finding has been reported by Rohrer (1949). In the face of such inconsistent evidence, Alper's (1948) contention that rehearsal plays a mediating role between reminiscence in rote learning and the variables that influence it, such as ego-involvement, is questionable. At any rate, rehearsal is entirely unsuitable as an explanation in the present case for

the following reasons: (1) The rest interval activity is engaging enough to give the subject no time for rehearsing. (2) Rehearsal calls for silent recitation of not only responses but also stimuli. In this sense, there could be little rehearsal, if any, in our experiment because of the use of non-verbal stimuli. (3) If rehearsal did occur and had a facilitative influence, the recall immediately after rest would not show the relative loss that it did. (4) Finally, rehearsal by no means can account for better evidence of reminiscence when ISS is high than when it is low.

Hull (1943) has offered a theory on the basis of his postulate of reactive inhibition I_R, which is sometimes invoked to explain reminiscence as well as distributed practice efficacy in both motor and verbal learning. According to this theory, every instance of responding, whether reinforced or non-reinforced, generates an inhibitory state which (1) is a positive function of the degree of effortfulness of response, (2) depresses performance temporarily, without directly interfering with the growth of habit strength, $_SH_R$, and (3) dissipates with rest periods. During a rest interval interpolated after a sufficient build-up of I_R by massed practice, both I_R and $_SH_R$ are dissipated as a mere result of lapse of time. However, I_R is dissipated more rapidly than $_SH_R$, and therefore the reaction-evocation potentiality of the stimulus increases in some degree through the rest interval.

The above theory fails to predict the phenomena of reminiscence in verbal learning in several ways: (1) As reactive inhibition continually accumulates with massed practice, greater reminiscence should occur when the degree of pre-rest learning is high than when it is low. This has not been found to be the case. The findings of Hovland (1938*a*, *b*, 1939), on the contrary, exhibit better evidence of reminiscence when rest is interpolated after moderate learning than after a high criterion has been attained. (2) Another implication of the reactive-inhibition theory is that if practice is considerably massed, that by itself may be a sufficient condition for reminiscence to appear. That this is not true has been proven in the experiments of Buxton (1949), Buxton & Bakan (1949), Buxton & Ross (1949), Melton & Stone (1942), Noble (1950), Shipley (1939) and Underwood (1957). In all these studies, practice was almost as massed as it could possibly be with a presentation rate of 2 sec. per syllable and an ITI of about 6 sec. This point has been proven particularly in the work of Melton & Stone (1942), who with a 6 sec. ITI failed to produce reminiscence even when the syllable presentation pace was increased from 2 to 1·45 sec. (3) As pointed out by Rohrer (1949), Hull's theory predicts a positive relationship between the amount of effort expended in learning and degree of reminiscence. The failure of this prediction is reflected in the lack of correlation found by Rohrer (1949) between the total number of responses (both correct and incorrect) elicited during acquisition and the amount of reminiscence exhibited during recall and by Champion & Scott (1953) between trials to pre-rest criterion and reminiscence scores. The prediction is similarly refuted by the product-moment correlation coefficients computed between trials to pre-rest criterion and post-rest trials to both 5/6 and 6/6 criteria in our experiment. The r's for the 0-, 2-, and 4-min. RPs are, respectively, 0·215, $-0·275$, $-0·083$ for the 5/6 criterion and 0·198, $-0·363$ and $-0·498$ for the 6/6 criterion. The tests of null hypothesis by Fisher's z' transformation method do not reveal a significant difference of r for 0-min. RP from r for either 2-min. or 4-min. RP in respect of either post-rest criterion.

Rohrer (1949) has advanced a theory of reminiscence in rote learning by postulating a temporally labile frustration-induced inhibitory state which is a positive function of the number of non-rewards during acquisition. Non-reward in this terminology means the occasion when the subject, soon after making a wrong response, becomes aware of it. The evidence in support of the theory which both Rohrer (1949) and Champion & Scott (1953) claim to have derived from certain experiments on serial learning is unacceptable for the following reasons: (1) Both experiments employed the double-recall method which has been proven to be a defective method of measuring reminiscence (see Gray, 1940). (2) Rohrer's (1949) conjecture that rehearsal or review causes added frustration is questionable. (3) The surmise of Champion & Scott (1953) that ego-involvement necessarily leads to frustration is of doubtful validity, for Dey & Kaur (1965) have shown that ego-involvement needs to be stepped up high enough before it may turn from a positively motivational state to a frustrating one.

Better evidence of reminiscence in PA learning was obtained with the HS than with the LS stimulus series in our experiment. It may be stated that the previous attempts to produce reminiscence in PA learning (Hovland, 1939; Riley, 1953, Expt. 2; Riley, 1954, Expt. 4) had failed because of the use of nonsense syllables as stimuli which presumably do not create a condition of sufficiently high ISS. A suitable explanation of reminiscence in PA learning should be broad-based enough to handle the dependency of this phenomenon upon ISS as well as the inverse relationship of ISS to the rate of PA learning. The latter relationship has been demonstrated in several previous studies (e.g. Dey, 1966*a*) and reconfirmed in the pre-rest data of the present experiment.

Gibson (1940) has formulated a theory of PA learning using certain fundamental principles of conditioning: stimulus generalization, differential reinforcement and spontaneous recovery. Gibson states that (1) during PA learning intra-list generalization occurs in the form of carry-over of responses to every stimulus from the other stimuli, (2) repeated non-reinforcement of the generalized (incorrect) responses leads to their extinction, and (3) the generalized responses which are thus extinguished may reappear after a lapse of time by the process of spontaneous recovery. With the assumption that the process of verbal associative learning is basically the same as that of discrimination learning, Gibson (1940) explains the inverse relationship of rate of PA learning to ISS. However, when it comes to the question of reminiscence, because of the contention that intra-list generalization spontaneously recovers from extinction, the Gibson theory would predict just the opposite effect—relatively greater forgetting—which would be a positive function of ISS.

Of all the assumptions used by Gibson (1940) in building her theory, the assumption that a process comparable to stimulus generalization in conditioning occurs in PA learning is the most viable. Solid support for this assumption is available from several experiments which prove that the empirical stimulus generalization gradient is as much valid with voluntary reactions as with conditioned response (cf. Dey, 1966*b*). Therefore the concept of intra-list generalization can still be used as an analytical device, though the Gibson theory seems ineffectual, and other avenues must be explored to speculate what implications intra-list generalization may possibly have for reminiscence. One such avenue would be to use the theory of differential forgetting (McGeoch, 1942) and argue that the wrong associations (generalized responses),

being relatively weak (Easley, 1937), are forgotten faster than the right associations (correct responses) through the rest interval, thereby showing a relative net gain in post-rest recall or extended learning. This would satisfactorily account for the relationship of reminiscence to ISS, since the benefit accruing from differential forgetting would conceivably be the greater the stronger the generalized responses. However, we have hesitation in accepting this explanation because of the proven vulnerability of the theory of differential forgetting (McClelland, 1942; Wilson, 1949).

A theory similarly framed to the one earlier advanced by the writer (Dey, 1966*a*) to account for distributed practice efficacy in PA learning may provide a satisfactory explanation of the reminiscence phenomena observed in the present experiment. Accepting Gibson's supposition that response is carried over by stimulus generalization from every component to every other component in a PA task, this theory assumes that in the performance of any component, as the appropriate response and the generalized responses are simultaneously evoked, the competition for overt elicitation generates an inhibitory state which is comparable to the competition-induced inhibition postulated by Wendt (1936), accumulates as the excitatory potential grows with repeated performance, and is temporally labile. The interpolation of a brief rest pause makes both excitatory and inhibitory potentials dissipate, but the dissipation rate being quicker for the latter, the effective excitatory potential (excitatory potential minus inhibitory potential) happens to be greater after the rest than if there was no such rest, thus exhibiting reminiscence. The theory further assumes that the magnitude of competition-induced inhibition is a positive function of the strength of the competing response which in turn bears a directly proportionate relationship to stimulus similarity. This additional assumption helps to explain both the inverse relationship of acquisition rate to ISS and the fact that sufficient ISS is a necessary condition for the occurrence of reminiscence in PA learning.

The theory offered by Hull *et al.* (1940) provides by far the best explanation of reminiscence in serial learning. The inhibition postulated in this theory is the inhibition of delay which results from a trace-conditioning-like process developed during the formation of serial associations. While successfully predicting the relationship of reminiscence to certain important variables, such as list length (Shipley, 1939) and serial position (Hovland, 1938*a*, *b*,) as well as certain salient facts concerning the influence of practice distribution on serial learning (Hovland, 1938*c*, 1940; Patten, 1938) and, above all, the serial position gradient itself, this theory fails to account for the absence of reminiscence when the serial items are words instead of nonsense syllables (Buxton, 1949; Melton & Stone, 1942; Noble, 1950). In our opinion, this defect in the theory of Hull *et al.* (1940) can be corrected if it is assumed that intra-list generalization occurs and, therefore, some amount of competition-induced inhibition is generated in serial learning, depending upon the similarity among list items, and that reminiscence in serial learning is the combined effect of the removal of both inhibitory potentials through rest pauses. Such a modification may enable the theory of Hull *et al.* (1940) to explain some further facts of relevance, such as the positive relationship of degree of facilitation of serial learning by practice distribution to similarity among list terms (Underwood & Goad, 1951; Underwood & Richardson, 1958; Underwood & Schulz, 1959).

The writer wishes to thank Mr Michael Jones and Miss Donna Van Alstine for their assistance in collecting data.

REFERENCES

ALPER, T. G. (1948). Task-orientation and ego-orientation as factors in reminiscence. *J. exp. Psychol.* **38**, 224–238.

ARCHER, E. J. (1953). Retention of serial nonsense syllables as a function of rest-interval responding rate and meaningfulness. *J. exp. Psychol.* **45**, 245–252.

BUXTON, C. E. (1949). Repetition of two basic experiments on reminiscence in serial learning. *J. exp. Psychol.* **39**, 676–682.

BUXTON, C. E. & BAKAN, M. B. (1949). Correction vs. non-correction learning techniques as related to reminiscence in serial anticipation learning. *J. exp. Psychol.* **39**, 338–341.

BUXTON, C. E. & ROSS, H. V. (1949). Relationship between reminiscence and type of learning technique in serial anticipation learning. *J. exp. Psychol.* **39**, 41–46.

CHAMPION, R. A. & SCOTT, E. (1953). Reminiscence and frustration-induced inhibition. *Aust. J. Psychol.* **5**, 1–9.

DEY, M. K. (1966*a*). An explanation of distributed practice efficacy in paired-associate learning. *J. gen. Psychol.* **74**, 61–75.

DEY, M. K. (1966*b*). Latency gradient of stimulus generalization with voluntary reactions. *J. gen. Psychol.* **74**, 289–298.

DEY, M. K. & KAUR, G. (1965). Facilitation of performance under experimentally induced ego-motivation. *J. gen. Psychol.* **73**, 237–247.

EASLEY, H. (1937). The curve of forgetting and distribution of practice. *J. educ. Psychol.* **28**, 474–478.

GIBSON, E. J. (1940). A systematic application of the concept of generalization and differentiation to verbal learning. *Psychol. Rev.* **37**, 196–229.

GRAY, S. (1940). The influence of methodology upon the measurement of reminiscence. *J. exp. Psychol.* **27**, 37–44.

HOVLAND, C. I. (1938*a*). Experimental studies in rote-learning. I. Reminiscence following learning by massed and distributed practice. *J. exp. Psychol.* **22**, 201–224.

HOVLAND, C. I. (1938*b*). Experimental studies in rote-learning theory. II. Reminiscence with varying speeds of syllable presentation. *J. exp. Psychol.* **22**, 338–353.

HOVLAND, C. I. (1938*c*). Experimental studies in rote-learning theory. III. Distribution of practice with varying speeds of syllable presentation. *J. exp. Psychol.* **23**, 172–190.

HOVLAND, C. I. (1939). Experimental studies in rote-learning theory. IV. Comparison of reminiscence in serial and paired associate learning *J. exp. Psychol.* **24**, 466–484.

HOVLAND, C. I. (1940). Experimental studies in rote-learning theory. VI. Comparison of retention following learning to the same criterion by massed and distributed practice. *J. exp. Psychol.* **26**, 568–587.

HOVLAND, C. I. & KURTZ, K. H. (1951). Experimental studies in rote-learning theory. IX. Influence of work-decrement factors on verbal learning. *J. exp. Psychol.* **42**, 265–272.

HULL, C. L. (1943). *Principles of Behavior*. New York: Appleton-Century-Crofts.

HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T. & FITCH, F. B. (1940). *Mathematico-deductive Theory of Rote Learning*. New Haven: Yale University Press.

IRION, A. L. (1949). Retention and warming-up effects in paired-associate learning. *J. exp. Psychol.* **39**, 669–675.

JUNG, J. (1966). A test of Underwood's theory of distributed practice. *J. exp. Psychol.* **71**, 778–780.

MCCLELLAND, D. C. (1942). Studies in serial verbal discrimination learning. I. Reminiscence with two speeds of pair presentation. *J. exp. Psychol.* **31**, 44–56.

MCGEOCH, J. A. (1942). *The Psychology of Human Learning*. New York: Longmans, Green.

MCGEOCH, J. A., MCKINNEY, F. & PETERS, H. N. (1937). Studies in retroactive inhibition, reproductive inhibition and reminiscence. *J. exp. Psychol.* **20**, 131–143.

MARSHALL, M. A. & RUNQUIST, W. N. (1962). Facilitation in paired-associate learning by distributed practice. *J. verb. Learn. verb. Behav.* **1**, 258–263.

MELTON, A. W. & STONE, A. R. (1942). The retention of serial lists of adjectives over short intervals with varying rates of presentation. *J. exp. Psychol.* **30**, 295–310.

Noble, C. E. (1950). Absence of reminiscence in the serial rote learning of adjectives. *J. exp. Psychol.* **40**, 622–631.

Patten, E. F. (1938). The influence of distribution of repetitions on certain rote learning phenomena. *J. Psychol.* **5**, 359–374.

Riley, D. A. (1953). Reminiscence effects in paired associate learning. *J. exp. Psychol.* **45**, 232–238.

Riley, D. A. (1954). Further studies of reminiscence effects with variations in stimulus response relationships. *J. exp. Psychol.* **48**, 101–105.

Riley, D. A. (1957). The influence of amount of pre-rest learning on reminiscence effects in paired associate learning. *J. exp. Psychol.* **54**, 8–14.

Rohrer, J. H. (1949). Factors influencing the occurrence of reminiscence: attempted formal rehearsal during the interpolated period. *J. exp. Psychol.* **39**, 484–491.

Shipley, W. C. (1939). The effect of a short rest pause on retention in rote series of different lengths. *J. gen. Psychol.* **21**, 99–117.

Underwood, B. J. (1957). Studies of distributed practice. XVI. Some evidence on the nature of the inhibition involved in massed learning of verbal materials. *J. exp. Psychol.* **54**, 139–143.

Underwood, B. J. (1961). Ten years of massed practice on distributed practice. *Psychol. Rev.* **68**, 229–247.

Underwood, B. J. & Ekstrand, B. R. (1967). Effects of distributed practice on paired-associate learning. *J. exp. Psychol. Monogr. Supp.* **73**, no. 4, pt. 2. (Whole no. 634.)

Underwood, B. J., Ekstrand, B. R. & Keppel, G. (1964) Studies of distributed practice. XXIII. Variations in response-term interference. *J. exp. Psychol.* **69**, 201–212.

Underwood, B. J. & Goad, D. (1951). Studies of distributed practice. I. The influence of intra-list similarity in serial learning. *J. exp. Psychol.* **42**, 125–134.

Underwood, B. J. & Richardson, J. (1958). Studies of distributed practice. XVII. The influence of meaningfulness and intralist similarity of serial nonsense lists. *J. exp. Psychol.* **56**, 213–219.

Underwood, B. J. & Schulz, R. (1959). Studies of distributed practice. XIX. The influence of intralist similarity with lists of low meaningfulness. *J. exp. Psychol.* **58**, 106–110.

Underwood, B. J. & Schulz, R. W. (1961). Studies of distributed practice. XXI. Effects of interference from language habits. *J. exp. Psychol.* **62**, 571–575.

Ward, L. B. (1937). Reminiscence and rote learning. *Psychol. Monogr.* **49**, no. 220.

Wendt, G. R. (1936). An interpretation of inhibition of conditioned reflexes as competition between reaction systems. *Psychol. Rev.* **43**, 258–281.

Whitey, S., Buxton, C. E. & Elkin, A. (1949). Control of rest interval activities in experiments on reminiscence in serial verbal learning. *J. exp. Psychol.* **39**. 173–176.

Wilson, J. T. (1949). The formation and retention of remote associations in rote learning. *J. exp. Psychol.* **39**, 830–838.

Winer, B. J. (1962). *Statistical Principles of Experimental Design*. New York: McGraw-Hill.

Woodworth, R. S. & Schlosberg, H. (1954). *Experimental Psychology*. New York: Holt, Rinehart & Winston.

(*Manuscript received* 1 *October* 1969; *revised manuscript received* 7 *January* 1970)

Br. J. Psychol. (1971), **62**, 1, *pp.* 53–58
Printed in Great Britain

MEMORY AS A RECONSTRUCTIVE PROCESS

By HOWARD R. POLLIO and RUSSELL FOOTE
Department of Psychology, University of Tennessee, Knoxville, Tenn., U.S.A.

This experiment investigated the effects of subject-defined structure on the free recall of verbal material. Four separate groups of 15 subjects each sorted four different word lists into individually defined categories. Immediately after sorting, the subjects were unexpectedly asked for a free recall of these words. All subjects' recall outputs were organized in concord with their sorting categories. Over all subjects a number of common categories emerged. In addition, recall categories were found to be temporally distinct in individual recall records. The results were taken to support the view that memory is a reconstructive rather than an associative process when task constraints do not preclude the subject from using subject-defined organizational factors.

The purpose of the present experiment was to investigate the effects of subject-defined organization on a number of different measures of free recall. In order to do this it was necessary first to provide subjects with an opportunity to organize materials in a concept-sorting task and then require them to recall these newly structured materials. In order to provide as stringent a test as possible subjects were not told during concept-sorting that they would later be required to recall the categorized items. The hypothesis of the present experiment was that each subject's own pattern of organization, as revealed in the groups of words sorted into a common category, would manifest itself in his pattern of recall: that is, subjects would tend during recall to produce items from the same category in adjacent or near adjacent positions in recall. In addition to these order effects it was expected that subjects would also exhibit shorter inter-response times (IRTs) for items from the same category when compared to IRTs separating items from different categories. This latter expectation is based on the results of a number of prior studies (e.g. Pollio *et al.*, 1968, 1969) in which it was found that the temporal properties of recall output are often correlated with the structure of the materials employed where such structure had been defined statistically in terms of word-association or category norms.

Method

The basic rationale of the present experiment required each subject first to impose his own structure on a set of verbal material in a sorting task, and then to produce an unexpected free recall of this material. The recall task was made an unexpected one in an attempt to prevent pre-recall rehearsal. The crucial point of the procedure, however, was to create an experimental condition in which it would be possible to analyse the effects of an individual's concept-structure on his subsequent pattern of free recall.

Subjects. Sixty undergraduate students who were enrolled in introductory psychology courses at the University of Tennessee were used. The median age of the sample was 19 and both sexes were about equally represented. Although all subjects were volunteers for this particular experiment, it was a part of each subject's course requirements that he participate in some experiment.

Stimuli. Four lists (22 words per list) differing in the number of inter-item associative connexions (see Pollio, 1968, for the procedure used to assess number of inter-item associative connexions in a word set) were used in this experiment, with the lists designated as MUSIC,

COMMAND, BUTTERFLY and RANDOM. Words in the MUSIC list were word associations to the stimulus word MUSIC, while words in both the COMMAND and BUTTERFLY lists were associates given in response to these two words, respectively. Words in the RANDOM list were chosen so that no word in the list evoked any other as a primary associate. These four different sets of words were used to ensure generality over materials. Associative norms indicate that the total number of inter-item associative connexions was 116 for the MUSIC set, 93 for the BUTTERFLY set, 73 for the COMMAND set, and 0 for the RANDOM set.

Each word in each of the four lists was typed on a vinyl card often used in the operation of the Hunter Card Master (Model 340). The specific words contained in each of the four lists are presented in Table 1.

Table 1. *The four sets of stimulus words*

Word set			
1. MUSIC	2. COMMAND	3. BUTTERFLY	4. RANDOM
Note	Demand	Yellow	Sword
Hear	General	Bee	Plain
Horn	Strong	Blue	Danger
Drum	Hear	Bug	Wise
Soft	Navy	Bird	Flower
Band	Stop	Wing	Bitter
Beautiful	Yell	Wasp	Money
Play	Order	Warm	Truth
Sing	Command	Butterfly	Glow
Nice	Attention	Cocoon	Joy
Music	Officer	Summer	Run
Opera	Force	Spring	Art
Pretty	Tell	Colour	Stars
Tune	Halt	Dog	Dawn
Sound	Do	Flower	Table
Tone	Talk	Sky	Boat
Voice	Go	Pretty	Silk
Piano	Hard	Flutter	Faith
Instrument	Shout	Moth	Nurse
Noise	Obey	Fly	Wagon
Song	Man	Light	City
Symphony	Army	Insect	High

Procedure. As each subject reported to the experimental room he was assigned to one of four groups, designated MUSIC, COMMAND, BUTTERFLY or RANDOM. All subjects were seated at a table in the experimental room and then presented with the appropriate set of 22 word-cards. The word-cards were placed before the subject in a stack, with different random orders for different subjects. Each subject was then given the following instructions:

'In front of you, as you can see, is a stack of white cards. On the top side of each card appears a word. What I would like you to do is sort these words into groups. I would like you to do this in the following manner: take the first card off of the stack, look at the word on it, and place the card on the table off to the side of the stack. This will expose the second card. If you think that the word appearing on the second card should be grouped with the word on the first card, place the second card with the first so that both cards remain exposed. If you do not think that the word which appears on the second card should be grouped with the word on the first, or if you are not sure, place the second card somewhere on the table away from the first, thus forming the beginning of another group. Then proceed similarly with the rest of the cards in the stack, taking cards off the stack one at a time. Place each card in the group where you think it belongs. If you think a card should not be grouped in any of the groups you have made so far, or if you are not sure, start a new group. Also, you may move cards from one group to another at any time if you decide that a word you previously placed in one group should, instead, be grouped differently....'

While the subject sorted the cards, the experimenter remained in the experimental room and answered any questions by saying 'Proceed in any way you like' or 'It's up to you'. (Very few questions were asked and it appeared that the task was a meaningful one for all subjects.)

Without the subject's knowledge, a record of total sorting time was obtained with a stopwatch, beginning with the removal of the first word-card from the top of the stack, and ending when the subject had completed his sort. At the conclusion of sorting, the subject was asked why he had sorted as he had, and, if possible, to suggest a word or phrase that might function as a title or descriptive label for each of the newly sorted groups. Finally, the subject was asked to rank these groups from first to last on the basis of how well the words within a group 'went together'. All one word groups were ignored in both the labelling and ranking tasks.

The subject was then asked to move to a chair in front of a microphone which was in another part of the experimental room and was instructed essentially as follows:

'What I would like you to do now is try to recall as many of the words you have just been working with as you can... You may recall them in any order you wish. You do not necessarily have to recall them in the same way you grouped them at the table. Recall them in whatever way that comes to mind...'

At the completion of the recall period, the subject was asked if he had expected—before being asked to move to the microphone—to recall the words. All subjects indicated they did not anticipate the recall task. Expressions of surprise were made by most subjects when they were asked to recall or when they were first exposed to the microphone.

While the subject was engaged in the recall task, the experimenter made a written record of the words in each of the categories produced. The location of the subject during his recall prevented him from looking at the cards. A Wollensack tape-recorder was used to record each subject's recall. The recorded recall was then used to actuate the pen of a single-channel Techni-Rite heat writing oscillograph (Model 4), with chart paper moving at a rate of 10 mm/min. After the recording had been made on chart paper, the tape was replayed to coordinate each deflection of the oscillograph pen with the word recalled by the subject. In this manner both a verbal and temporal record was obtained for each subject's recall.

Results

Recall means and standard deviations. The means and standard deviations for all words recalled (tokens), for all words recalled excluding repeats and intrusions (types), for all repeats, and for all intrusions are presented separately for each stimulus list in Table 2. To test for the possible effects of stimulus lists upon these characteristics of recall, analyses of variance were computed. Results of these analyses failed to reveal significant list effects for tokens (F = 2·21; d.f. = 3, 56), repeats (F = 1·04; d.f. = 3, 56), intrusions (F = 1·85; d.f. = 3, 56) or types (F = 1·64; d.f. = 3, 56). As may be seen, under present conditions, randomly selected words are recalled at least as well as words selected to exhibit strong associative interconnectedness.

Table 2. *Recall data for the four different word lists*

Recall variables	Stimulus word lists			
	MUSIC	COMMAND	BUTTERFLY	RANDOM
Tokens recalled				
Mean	17·6	18·7	20·0	18·6
S.D.	3·20	3·23	2·92	2·27
Repeats				
Mean	1·00	1·13	1·00	0·40
S.D.	0·73	1·46	1·65	0·61
Intrusions				
Mean	0·40	0·27	0·13	0·07
S.D.	0·61	0·45	0·34	0·25
Types recalled				
Mean	16·2	17·3	18·9	18·1
S.D.	2·66	3·19	2·38	2·68

Inter- and intra-category IRT. Since previous work (Pollio *et al.*, 1968, 1969) has shown that the temporal characteristics of recall output are frequently indicative of recall organization, an analysis was made of IRTs separating words in recall. Two types of IRTs were defined: intra- and inter-category IRTs. An intra-category IRT was defined as the time elapsed between the offset of one word in a category and the onset of the next word from that same category. Inter-category IRT was defined as the time elapsed between the offset of the last word in one category and the onset of the first word in the next category. Recall categories for each subject were defined in terms of each subject's own categories in sorting.

Table 3 presents the mean intra- and inter-category IRTs for all four lists combined, as well as the mean values for each stimulus list separately. In each line of the table, the first five values specify the mean between-category IRT for the first six categories in recall while the remaining six values specify the mean within-category IRT for latencies between sequential items in categories 1 to 6. In the case of these within-category IRTs, all values present a mean computed over all IRTs separating words from the same category. As before, these categories were defined in terms of categories produced by each subject during sorting. These results clearly indicate that for each stimulus list and for all word lists combined mean IRTs between words from the same category are faster and less likely to increase sharply in later recall than is true for IRTs between words from different categories.

Table 3. *Mean IRT between and within recall categories*

	Mean IRT (sec.)										
	Between sequential categories in recall					Between sequential items within categories in recall					
Word set	1	2	3	4	5	1	2	3	4	5	6
MUSIC	1·20	3·67	3·71	3·35	5·61	0·87	1·14	1·84	2·83	1·94	0·87
COMMAND	1·09	1·41	2·27	5·88	10·58	0·67	0·58	0·97	1·03	2·02	0·91
BUTTERFLY	2·05	1·83	2·19	2·47	5·65	0·45	0·92	0·37	0·59	1·17	1·70
RANDOM	1·09	1·95	1·95	4·84	4·91	0·43	0·54	0·93	0·90	1·84	0·62
All lists combined	1·36	2·22	2·53	4·00	6·55	0·69	1·12	1·35	1·40	1·90	1·15

In an attempt to assess the significance of such trends a three-factor analysis of variance was performed on these IRT data with one factor, word list; a second factor, order of output; and a third factor, type of IRT (within- or between-category). This analysis was confined to IRTs occurring for the first five recall categories only as the median number of recall categories for all subjects equalled three. Since there was no *a priori* basis for comparing each subject's first inter-category IRT with either the first or second intra-category IRT, the mean of these two IRTs for categories 1 and 2 was used. Similarly, the mean of the intra-category IRTs of categories 2 and 3 was compared with the second inter-category IRT, and so on for categories 3 and 4 and 4 and 5. This resulted in an analysis of variance involving four intra-category and four inter-category IRTs. Since 22 subjects in the experiment had either one, or more, one-word categories in recall (thus, not permitting the computation of an intra-category IRT) or did not produce at least five recall categories, a total of 30 observations was missing from the total of 480 observations called for in the design. These

missing cell values were estimated by using the mean value of that output order position. Degrees of freedom were subtracted from appropriate error terms.

The results of this analysis of variance revealed no effect of differences in list ($F = 1{\cdot}34$; d.f. $= 3, 56$), although a significant IRT-type effect, within *v.* between ($F = 32{\cdot}77$; d.f. $= 1, 56$; $P < 0{\cdot}01$) and a significant output order effect ($F = 6{\cdot}97$; d.f. $= 3, 166$; $P < 0{\cdot}01$) were in evidence. In addition, a significant interaction of IRT-type with output order ($F = 3{\cdot}21$; d.f. $= 3, 140$; $P < 0{\cdot}05$) was also found. All other interactions yield F values less than unity. From this analysis, IRTs occurring between items from the same category are significantly shorter than those that occur between items from different categories, with these inter-category IRTs increasing over later output positions. The significant interaction of IRT type with output order suggests that inter-category IRTs increase regularly, while intra-category IRTs increase irregularly if at all.

Discussion

In terms of memory, then, the present results indicate that how the subject sorts words into meaningful clusters will strongly determine the order and temporal properties of his output protocol. The match-up between sorting and recall data both in terms of pattern and temporal properties supports this conclusion directly. Subjects, under incidental learning instructions, tend to produce words during recall in terms of their own prior sortings, and tend to pause in their output when shifting from one category to another. As in previous experiments, this increase in between-category IRTs is taken to reflect the time required for subjects to retrieve categories (not items) in memory storage, with later categories taking longer to retrieve than earlier ones. The pattern of within-category IRTs indicates that the individual items, once categorized, essentially function as a single unit in memory with only slight increases in reconstructing later categories relative to earlier ones.

All in all, present results suggest that recall of categorized materials is composed of two components the first of which involves retrieval of the category itself and the second of which involves the production of individual items. Although in a previous paper (Pollio & Gerow, 1968) we suggested that individual items were reconstructed during the process of recall itself, the present results do not unequivocally rule out other possibilities. For example, present results might be taken to support a word-to-word linkage theory of the following sort: each subject orders his cards in such a way during sorting that he presents items to himself as stimuli in clusters, and then later recalls these items in accordance with the linkages established during sorting.

This analysis, however, runs into some difficulty in terms of the procedures used in the present study. For one, the subjects did not expect a recall test and therefore there is no reason to assume they paid any attention to item order in their initial sorting behaviour. Secondly, the instructions clearly allowed the subjects to 'move cards from one group to another' and most subjects did in fact move items around, making stable linkages difficult to achieve. Finally, an examination of the terminal order achieved in any given category in sorting failed to predict systematically the order of items produced in recall. All in all, these arguments suggest that the subjects did not establish many (or any) item-to-item linkages in sorting, and that some sort of regeneration of items probably occurred during recall itself.

According to this type of analysis, once a category is recalled, items are regenerated (or reconstructed) on the basis of just those attributes causing them to be sorted together in a single category in the first place. Obviously, those words serving to define a particular category best are precisely those words best recalled. Given this consideration, it is unlikely that subjects would be able to generate all of the items in any given category, since words less-defining of the category would be difficult if not impossible to regenerate. This would be particularly true in the present case since all items were seen only once during an incidental learning task.

This research was supported by Grant MH 80903 from the United States Public Health Service. A preliminary report of this work was presented at the Experimental Psychology Society Meetings in Oxford, England, in March 1969. The present paper was written while the senior author held a Special Research Fellowship from the USPHS and was in residence at the Psycholinguistics Research Unit, Department of Phonetics, University of London.

REFERENCES

POLLIO, H. R. (1968). Associative structure and verbal behavior. In T. R. Dixon & D. L. Horton (eds.), *Verbal Behavior and General Behavior Theory*. Englewood Cliffs, N.J.: Prentice-Hall.

POLLIO, H. R. & GEROW, J. R. (1968). The role of rules in recall. *Am. J. Psychol.* **81**, 303–313.

POLLIO, H. R., KASSCHAU, R. A. & DENISE, H. E. (1968). Associative structure and the temporal characteristics of free recall. *J. exp. Psychol.* **76**, 190–197.

POLLIO, H. R., RICHARDS, S. & LUCAS, R. (1969). Temporal properties of category recall. *J. verb. Learn. verb. Behav.* **8**, 529–536.

(*Manuscript received* 4 *October* 1969; *revised manuscript received* 26 *February* 1970)

Br. J. Psychol. (1971), 62, 1, pp. 59–65
Printed in Great Britain

FREE RECALL WITH ASSISTANCE FROM ONE AND FROM TWO RETRIEVAL CUES

BY P. D. McLEOD,* C. E. WILLIAMS† AND D. E. BROADBENT*
Psychological Laboratory, University of Cambridge

A list of words was presented for free recall, and hints were later given for unrecalled words by presenting first one and then a second cue word associated with the unrecalled words. Earlier results by Tulving & Osler had shown that one cue assists recall but had shown no benefit from the second cue; in their task, however, the associations were established with the experiment. In the present results, the second cue did produce a further increment in recall; furthermore, the increment in recall was greater than that predicted on the assumption that there was no interaction between the two cues. This means that the first cue had produced some effect even when it did not give full recall.

Tulving & Osler (1968) investigated the effect of single and double retrieval cues in aiding recall of a list of words. The cues they used were words to which the associated stimulus words occurred as a free-association response with $P < 0{\cdot}01$. In the condition of interest they compared recall with one cue word and with two cue words; in each case the word or words had been presented with the stimulus word during learning and the same words presented as cues in recall. They found that subjects recalled no more words with two cues to help them than with one. But this rather surprising result may have been due to two effects. First, in the stimulus plus two cue learning condition 25 sets of three words were presented for only 5 sec. each. This may have been insufficient time for learning to take place, in which case the 'cues' were not performing that function, since it depended on association during the experiment. Secondly, the two cues may have been stored as a single unit, and so were effectively one cue and not two.

Tulving & Osler found in another condition that if two cues were presented with the stimulus word in learning, and then either one or both were presented to aid recall, the subject did better with two than with one. However, there are three possible confusion effects which prevent one from interpreting these results conclusively as a demonstration of an interaction between two cues: (1) presentation of single cues in recall may result in recall of the second cue and not of the stimulus word; (2) presentation of one cue might lead to recall of the stimulus word via the second cue, making recall with one cue and two cues effectively identical; (3) if, as suggested before, the two cues are stored as one unit, presentation of only one may hinder retrieval of the stimulus word.

The present experiment was designed to study the possible interaction between two cues, i.e. to see whether two cues presented together facilitate recall to a greater extent than either would when presented separately, without the possible confusions of the Tulving & Osler design.

The topic is of some importance because it allows some separation of theories of retrieval depending on whether the difficulty of recall is due to failure to find certain words altogether, or to their low strength when found.

* Present address: MRC Applied Psychology Unit, Cambridge.
† Present address: Department of Experimental Psychology, Oxford.

Main experiment

Method

Subjects. Thirty-four undergraduates were used.

Material. The stimuli consisted of 25 nouns or adjectives in common English usage. The words were taken from the free-association norms of Bilodeau & Howell (1965) such that the stimuli occurred as free-association responses to their two cues with as high a probability as possible; thus any facilitation of retrieval produced by the cues would not depend on learning during the experiment. They were chosen with the restraining condition that the two cues should not occur as free-association responses to each other with $P > 0{\cdot}01$.

Procedure. All instructions were tape-recorded. The stimuli were presented one at a time on cards for 5 sec. After this a mental arithmetic task was given to reduce serial-position effects (Glanzer & Cunitz, 1966); this took most subjects about 20–30 sec. Free recall of the stimulus words was then asked for. Subjects were told not to guess, but only to say a word if they were more or less certain it was in the original list. After 2 min. subjects were told they would be given some help to recall those words which they had failed to recall so far. They were told the relationship between the cues and the stimulus words. For each unrecalled word a card giving one cue was placed in front of the subject. The instruction on guessing was repeated. Two min. were allowed for recall. The subject was then told that he would be given further help, with a second cue, and for each unrecalled stimulus word a card containing the previous cue (which had failed to produce recall) and a new cue was presented. Half the subjects received one cue first and the remainder received the other.

Practice. Prior to the experiment a practice run was held which required free recall of a list of 25 words presented for 5 sec. each. It was followed by a mental arithmetic task comparable in difficulty to the one used in the main test. Subjects were told what free association is and given practice at the rapid generation of free-association responses.

Results

Individual scores are shown in Table 2 after the Appendix, where they are more relevant.

The average probabilities of word recall were: with no cues, $PN = 0{\cdot}59$; after one cue, $PC_1 = 0{\cdot}73$; after two cues, $PC_2 = 0{\cdot}84$. (In the practice list the average recall was: PN practice $= 0{\cdot}55$.) The performance of every subject except two was better after two cues than one.

Control experiment

A defect with this experiment appeared to be the possibility that, in spite of the instructions, subjects were guessing the stimulus words when they saw the cues, and not remembering them. In particular, it might be supposed that with two cues generating dissimilar lists of words by free association, a word common to the two lists was likely to be the stimulus word. Accordingly, a control experiment was carried out using 20 undergraduates who had not taken part in the original experiment. They were told the relationship between cue words and stimulus words, and given the same practice at free associating as the subjects in the first experiment. They were then given the individual cues and then the double cues and asked to guess what the stimulus words were. They gave their first choice, and then any other words they thought might be likely. They were given the same length of time to guess as the average exposure time to the cues of the subjects in the main experiment: 12 sec. for one cue and 18 sec. for two cues.

It can be seen from Table 1 that guessing could make a small contribution to correct recall. The surprising result is that significantly fewer words were guessed with two cues than with one ($P < 0{\cdot}01$). The control experiment indicates that the apparent effect due to interaction of two cues will be reduced rather than enhanced by guessing.

A further point which suggests that little guessing was involved in the main experiment was the very low incidence of incorrect responses—none, in fact, for the majority of subjects.

This control experiment raises the problem of the mechanism by which the cues assist retrieval. It was originally supposed that the free-association responses to the cues would contain the original stimulus words, and these would be recognized by the subject as belonging to the original list. But in the control experiment the percentage of stimulus words guessed is too low to account for the number of stimulus words produced in the original experiment. It is possible that the presence of the stimulus word in memory, due to its presentation in the earlier part of the experiment, increases the probability of its being produced as a free-association response to the appropriate cue.

Table 1. *Average probabilities of correct responses with one cue and with two cues in the control experiment*

	One cue		Two cues	
	1st response	All responses	1st response	All responses
P(correct)	0·17	0·24	0·14	0·16

Mathematical models and results

Since the second cue does seem to improve recall, the results of the first experiment were compared with the predictions of three models, Additive, Power and Strength, which are derived in the Appendix. Briefly, the additive model argues that a cue which has been given, but failed to produce recall, leaves the system unaffected. Thus the effect of the second cue in such cases is the same as if no first cue had been given. The other models represent different ways in which the first cue might enhance the effect of a second. The obtained value of PC_2 was compared with that predicted by each model given the observed values of PC_1 and PN.

It was found that the value of PC_2 predicted by the additive model was always less than or equal to that predicted by the strength model and both values were always less than that of the power model prediction. Of the 34 subjects, nine performed less well than would be predicted by the additive model, two between additive and strength, 11 between strength and power, and 12 better than power (see Table 2).

The averages of PC_2 predictions were: additive, 0·824; strength, 0·837; power, 0·904; and observed, 0·844.

A Wilcoxon matched-pairs signed-ranks test applied to the individual data showed that: (i) subjects did better than additive model prediction with $P < 0{\cdot}002$; (ii) subjects did better than strength model prediction with $P < 0{\cdot}03$; and (iii) subjects did worse than power model prediction with $P < 0{\cdot}003$.

Conclusion

The main criticism of the Tulving & Osler design is the possibility that since, in conditions where two cues were to be associated with a single stimulus, the two cues were always presented together, the two cues were stored as a unit and so were effectively one cue. The experimental design does not allow for a separation of this effect from any possible interaction between the cues produced solely by their independent associations with the stimulus word.

The present experiment eliminates this possible confusion by the use of two cues which do not associate to each other. The results indicate that subjects are doing significantly better than would be predicted by the additive model. This shows that there is a positive interaction between the cues. The results cannot really be interpreted as favouring either of the other models put forward.

If there is a positive interaction between retrieval cues, then the help given by them in recall is not 'all or none', but continuous. Thus two cues, by themselves insufficient to produce a response, may do so when presented together.

There is the possibility that different types of cues operate in fundamentally different ways. It would be informative to repeat the experiment using the Tulving & Osler learning paradigm, but with non-simultaneous presentation of the two cues, thus establishing independent associations between the two cues and the stimulus word.

This experiment was performed as part of the requirements for Part II of the Natural Sciences Tripos (Psychology) at Cambridge University. It was carried out by the first two authors and supervised by the third.

References

Bilodeau, E. A. & Howell, D. C. (1965). *Free Association Norms by Discrete and Continued Methods.* Washington: Office of Naval Research Department of Navy.

Glanzer, M. & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *J. verb. Learn. verb. Behav.* **5**, 351–360.

Morton, J. (1969). Interaction of information in word recognition. *Psychol. Rev.* **76**, 165–178.

Tulving, E. & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *J. exp. Psychol.* **77**, 593–601.

(*Manuscript received* 12 *December* 1969; *revised manuscript received* 18 *February* 1970)

APPENDIX

Additive model

Conceptual background. This model assumes that there is no positive interaction between two cues in aiding retrieval. It supposes that each cue is attached to the correct response with a certain probability C, i.e. the cue will elicit the correct response with a probability C. In addition, a proportion PN of correct responses is available even in the absence of cues. The key assumption for additivity is that a cue, presented after a preceding cue, still has the same probability C of eliciting any correct response amongst those which have not already been given, i.e. attachment of one

cue does not make attachment of another cue more or less likely, and the numbers of responses obtained in each way are simply added together.

Another way of expressing this lack of interaction is: if each cue presented separately is ineffective in aiding recall, then no additional help will be obtained if the two cues are presented together.

Formulae. Given the original definitions of PN, PC_1, PC_2, on the theory

$$PC_1 = PN+(1-PN)C. \quad (1)$$

$$PC_2 = PC_1+(1-PC_1)C. \quad (2)$$

From eqn. (1)

$$\begin{aligned} 1-PC_1 &= 1-PN-(1-PN)C \\ &= (1-C)(1-PN). \end{aligned}$$

From eqn. (2)

$$\begin{aligned} 1-PC_2 &= (1-C)(1-PC_1) \\ &= (1-C)^2(1-PN) \\ &= \frac{(1-PC_1)^2}{(1-PN)^2}(1-PN) \\ &= \frac{(1-PC_1)^2}{1-PN}. \end{aligned}$$

Therefore

$$PC_2 = 1-\frac{(1-PC_1)^2}{1-PN}.$$

Strength model

Conceptual background. The strength model is a logogen model (Morton, 1969). It supposes that each possible response has a number attached to it, which is increased by the presence of the word in memory, by the presentation of one cue, *and* by that of a second one. The probability of a response is some function of the final size of the number. Thus the number of responses obtained in each condition should not simply add together, since, when one cue is given, the fact of earlier presentation of another cue should have increased the strength of some correct responses, even though insufficiently to produce overt response.

Formulae. For each response suppose a strength V, such that probability of that response $= V/\Sigma V$.

Let the strength of a response for a stimulus which was presented $= N$ in the absence of a cue. Let the sum of all *other* responses $= T$.

Then

$$PN = N/(N+T).$$

or

$$PN/(1-PN) = N/T.$$

When one cue is presented, the strength of the response corresponding to this cue is multiplied by C to give CN, the new strength.

It is assumed that when two cues are presented, the response strength is C^2N. Then

$$PC_1/(1-PC_1) = CN/T,$$

$$PC_2/(1-PC_2) = C^2N/T.$$

Therefore

$$PC_2/(1-PC_2) = [PC_1/(1-PC_1)]^2 \times (1-PN)/PN.$$

Therefore

$$PC_2 = \frac{[PC_1/(1-PC_1)]^2 \times (1-PN)/PN}{1+[PC_1/(1-PC_1)]^2 \times (1-PN)/PN}.$$

Power model

Conceptual background. This model supposes that each cue eliminates a certain number of words that were still possible following the earlier cue, and that the probability of a correct response is given by

$$P = \frac{1}{\text{no. of words still possible}}.$$

It is supposed that each cue eliminates the same *proportion* of words still in play, even if an earlier one has already eliminated some. The resulting effects on probability of correct response from presentation of two cues separately do not simply add because each cue is acting on a different population of words.

Formulae.

$$PN = 1/\text{original number} = 1/S.$$

$$C = \text{amount by which presentation of a cue divides the number of possibilities.}$$

Therefore

$$PC_1 = C/S = CPN.$$

$$PC_2 = C(PC_1)$$

$$= C^2PN.$$

Therefore

$$PC_2 = (PC_1)^2/PN.$$

Table 2 *Observed results for 34 subjects and the value of PC_2 predicted for each by the three models described in the text*

Subject	Observed values			Predicted values of PC_2		
	PN	PC_1	PC_2	Additive	Strength	Power
1	0·40	0·48	0·56	0·550	[illegible]	0·576
2	0·72	0·88	0·92	0·950	0·954	1·07
3	0·80	0·88	0·88	0·930	0·930	0·951
4	0·60	0·80	0·92	0·900	0·914	1·06
5	0·60	0·76	0·96	0·855	0·869	0·955
6	0·76	0·92	0·96	0·975	0·975	1·11
7	0·56	0·64	0·80	0·705	0·714	0·728
8	0·56	0·64	0·80	0·705	0·714	0·728
9	0·48	0·68	0·80	0·800	0·826	0·969
10	0·64	0·84	0·96	0·928	0·936	1·08
11	0·68	0·80	0·96	0·875	0·883	0·945
12	0·40	0·64	0·76	0·783	0·822	1·06
13	0·56	0·68	0·88	0·769	0·786	0·834
14	0·80	0·88	0·96	0·930	0·930	0·954
15	0·40	0·56	0·72	0·677	0·706	0·785
16	0·72	0·84	0·92	0·907	0·910	0·972
17	0·64	0·84	0·92	0·928	0·936	1·08
18	0·68	0·80	0·96	0·875	0·892	0·945
19	0·52	0·60	0·72	0·667	0·680	0·686
20	0·60	0·64	0·80	0·675	0·682	0·684
21	0·60	0·68	0·84	0·745	0·746	0·780
22	0·80	0·88	0·96	0·930	0·934	0·951
23	0·72	0·84	0·96	0·907	0·910	0·972
24	0·32	0·44	0·64	0·539	0·577	0·609
25	0·68	0·84	0·92	0·919	0·930	1·05
26	0·60	0·80	0·92	0·900	0·914	1·06
27	0·72	0·84	1·00	0·907	0·912	0·972
28	0·72	0·96	0·96	0·993	1·00	1·28
29	0·32	0·44	0·48	0·539	0·577	0·609
30	0·52	0·68	0·84	0·788	0·840	0·880
31	0·60	0·80	0·84	0·900	0·914	1·06
32	0·60	0·76	0·92	0·855	0·869	0·955
33	0·60	0·72	0·76	0·805	0·812	0·865
34	0·24	0·40	0·64	0·525	0·581	0·663

Br. J. Psychol. (1971), **62**, 1, pp. 67–72
Printed in Great Britain

IMAGERY AND SEX DIFFERENCES IN INCIDENTAL RECALL

By CAROLE H. ERNEST and ALLAN PAIVIO
Department of Psychology, University of Western Ontario, London, Canada

Male and female subjects who differed on measures of imagery ability were tested for incidental recall in two experiments involving pictures and words as stimuli. In one experiment high-imagery males surpassed their low-imagery counterparts in incidental free recall of words but the reverse relation occurred with females. No relation was obtained between imagery ability and incidental recall for the picture items. However, the stimuli were recalled better when associated with pictures rather than words as stimuli, suggesting an effect of stimulus concreteness on visual memory. The orientating task in the second experiment was item recognition. High-imagery subjects were more accurate than low-imagers in the recognition task, and high-imagery females (but not males) surpassed their low-imagery counterparts in incidental recall of the stimuli.

The significance of visual imagery as a factor in learning and memory has been confirmed in numerous studies (see Paivio, 1969, for a review). Recent investigations have suggested a relationship between individual differences in imagery and incidental learning, thus extending the range of learning tasks for which imagery is a relevant variable. Sheehan & Neisser (1969), for example, found that ratings of the vividness of imagery and accuracy in recalling incidentally presented materials were highly correlated, and Ernest & Paivio (1969) reported a positive relationship between imagery ability and incidental learning in a paired associates (PA) task where the stimulus or response terms were colour word compounds. The primary purpose of the present investigation was to test the generality of the latter finding. The first of two experiments involved a free-recall task (Type II incidental learning design) and the second experiment involved item recognition (Type I design). It was expected that subjects with high imagery ability, as measured by tests of spatial ability, would be superior to low-imagers in both experiments when presented with an incidental recall task.

The second purpose of this study was to explore the relationship between imagery ability and sex differences—a relationship essentially overlooked in recent investigations (e.g. Christiansen, 1969; Stewart, 1965). Ernest (1968) found that these two factors significantly interacted in a PA learning task such that differences between high- and low-imagers were greater for females than males. Moreover, high imagery ability was positively related to learning in the female sample, whereas the converse was found with males. A replication of this finding in other learning situations would have direct implications in the design and interpretation of studies incorporating imagery ability as a factor.

Experiment I

Method

Subjects. Four hundred and one introductory psychology students completed an Imagery Test Battery (see Ernest & Paivio, 1969) consisting of the revised Minnesota Paper Form Board (MPFB), the MPFB Questionnaire, and Space Relations. Subjects were defined as high, medium, or low imagers on the basis of the sum of their standard scores on each of the three tests. Sixty-six subjects participated in the present experiment, with nine males and 13 females in each

imagery level. The mean total standard scores for high, medium and low imagery were +3·016, +0·005 and −3·022, respectively. Verbal associational ability was controlled across imagery levels and sex using 60 sec. total associative production scores to eight stimulus nouns as a criterion.

Stimulus materials. Eighteen words were selected from a word pool of 925 nouns which had been scaled for abstractness–concreteness (C), image-evoking capacity (I), and meaningfulness (m) (Paivio *et al.*, 1968). Words were chosen which had high I and C ratings (i.e. their average ratings are greater than 6 on a 7-point scale) and whose pictorial referents, according to three judges, were not strongly associated with one specific colour. The I values for the 18 words varied from 6·17 to 6·87 (M = 6·48); the C values from 6·38 to 7·00 (M = 6·91). The m scores for the selected words varied from 5·68 to 7·88, with a mean of 6·88.

The words were printed in black outline upper-case letters and their pictorial representations were illustrated as black outline drawings. Each word and its referent picture were painted in one of nine oil pastel colours. The item–colour combinations were: arrow–brown, bird–red, book–black, bottle–orange, bowl–pink, butterfly–orange, car–black, clock–white, doll–green, dress–pink, flag–blue, kettle–white, pencil–brown, queen–red, table–green, umbrella–yellow, vest–blue and wigwam–yellow.

Two 18-item mixed lists of nine words and nine pictures were constructed with the restriction that a word and its pictorial representation not appear in the same list and each colour occur twice, once embedded in a word and once in a picture. Lists 1 and 2 differed in that the words and pictures of List 1 became the pictures and words, respectively, of List 2.

Procedure. The subjects were tested in groups in a darkened room. Each of the three imagery levels and both sexes were represented at each testing. The free-recall instructions, although informing the subjects that the items were painted in different colours, stressed the recall of the words and pictures only. The items were presented as 35 mm slides at a 2·25 sec. rate, using an Anscomatic II slide projector equipped with a 500 W. Proximity Reflector lamp. The rate included an exposure time of 1·25 sec. per item plus a 1 sec. blank period resulting from the automatic slide change. Two minutes were allowed for written recall. The light from two 25 W. overhead lamps (which remained on during the experiment) provided adequate lighting for subjects to record their responses. Two trials were presented, with the items in a different order each time. After Trial 2, a test of incidental learning was given. The subjects were presented all the items as black outlines only and instructed to write down the colour in which each item had been painted. Items were presented at 8 sec. intervals. Finally, the coloured test items were again presented, at 8 sec. intervals, and subjects recorded the names they had used to identify the pictures in the free recall task as well as the colour in which each item was painted. This procedure provided a basis for the objective scoring of each subject's free recall and incidental learning data.

Results

Free recall. Preliminary analyses revealed no recall differences between List 1 and List 2, so the data were collapsed across lists in subsequent analyses. An analysis of variance (Winer, 1962) incorporated imagery (high, medium and low), sex, stimulus attribute (pictures and words), and trials as factors, the latter two being within-subject factors. Significant main effects were obtained for sex (F = 7·31; d.f. = 1, 60; P < 0·01), attribute (F = 28·43; d.f. = 1, 60; P < 0·001) and trials (F = 261·24; d.f. = 1, 60; P < 0·001). Females recalled more items (M = 6·87) than males (M = 6·30), pictures were recalled better than words (7·11 *v.* 6·17), and recall improved from Trial 1 to 2 (5·73 *v.* 7·55). A significant two-way interaction (attribute × trials; F = 13·05; d.f. = 1, 60; P < 0·001) revealed that recall differences between pictures and words were greater at Trial 1. The significant triple interaction involving imagery, sex and attribute (F = 3·59; d.f. = 2, 60; P < 0·05) is illustrated in Fig. 1. High-imagery males consistently recalled more items than low-imagery males, the difference being greatest when words served as stimuli. A crossover between picture

and word scores occurred with females, low-imagers recalling more words than high-imagers.

Incidental learning. Since the correlation between total free-recall scores and total incidental learning scores was not significant ($r = 0{\cdot}10$), incidental recall scores were analysed using a $3 \times 2 \times 2$ analysis of variance (rather than covariance) with imagery, sex and stimulus attribute as factors. Only stimulus attribute emerged as a significant effect ($F = 153{\cdot}59$; d.f. $= 1, 60$; $P < 0{\cdot}001$). The colours of pictures were recalled significantly better ($M = 5{\cdot}14$) than the colours of words ($M = 2{\cdot}35$).

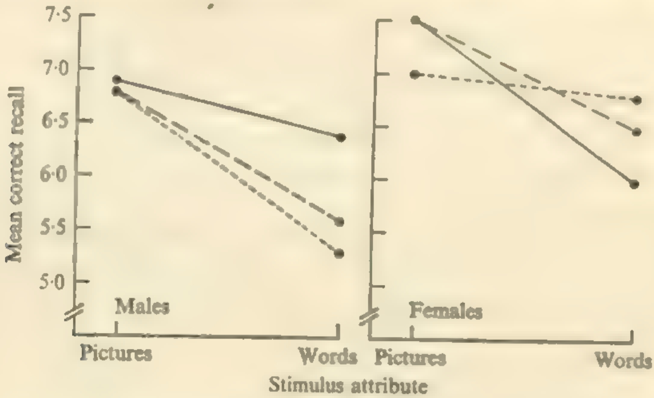


Fig. 1. Free recall as a function of imagery ability, sex and stimulus attribute. ——, High imagery; – – –, medium; ----, low.

Experiment II

Method

Subjects. Seventy-two subjects were selected from the pool of 401 subjects described in Expt. I. None had participated in the previous experiment. Thirty-six subjects were high-imagery and 36 were low-imagery. Males and females were equally represented. The mean total standard score was $+2{\cdot}274$ for the high- and $-2{\cdot}434$ for the low-imagery subjects. Verbal associational ability was controlled across sex and imagery level as before.

Stimulus materials. Three homogeneous lists of 72 pictures, their concrete noun labels, and 72 abstract words were constructed, using items from the pool of 925 nouns (Paivio *et al.*, 1968). The concrete and abstract nouns selected differed in ratings of C (6·87 and 2·54) as well as I (6·55 and 3·26). Meaningfulness and frequency (Thorndike & Lorge, 1944) were held relatively constant across levels of C and I. The m values for concrete and abstract words were 6·85 and 5·64 respectively; the mean frequencies were 46·26 and 43·25 occurrences per million, respectively.

Procedure. At each group testing session one of the three stimulus lists was presented. An attempt was made to have equal numbers of male and female subjects of each imagery level at each session. Testing terminated when six males and six females from each imagery level were tested for each of the picture, concrete word, and abstract word lists.

The stimuli were presented as 35 mm white-on-black-background slides projected on a screen. A motor-driven rotary disk shutter connected to the projector activated a slide change every 5 sec. An open sector in the disk allowed each slide to be projected for $\frac{1}{16}$ sec. In the remaining $4\frac{15}{16}$ sec. the subjects recorded the name of the item exposed. Three practice items were presented prior to the test list. The subjects were encouraged to guess.

The instructions emphasized the task as one of item recognition. After all items had been presented once, and the answer sheets collected, a measure of incidental learning was taken. The subjects were allowed 8 min. to record, in any order, as many of the 72 items as they could remember.

A scoring system unique to the picture list was devised, since subjects tended to give alternative labels to some of the pictures. Two judges (the authors) decided which alternative responses were acceptable. These responses were typically synonyms of the expected responses. For example, 'bugle' and 'horn' were accepted in lieu of 'trumpet'. In the incidental learning task, subjects were given credit for the correct recall of *any* item recorded in the item recognition task regardless of whether the item had been judged as acceptable or unacceptable. This scoring system precluded biasing incidental learning scores as a result of inaccurate recognition and was also used in the scoring of the concrete and abstract word lists.

Results

Item recognition. An analysis of variance revealed significant effects for imagery ($F = 10{\cdot}01$; d.f. = 1, 60; $P < 0{\cdot}01$), stimulus attribute ($F = 27{\cdot}83$; d.f. = 2, 60; $P < 0{\cdot}001$), and imagery by stimulus attribute ($F = 3{\cdot}40$; d.f. = 2, 60; $P < 0{\cdot}05$). High-imagery subjects recognized significantly more items ($M = 70{\cdot}72$) than low-imagery subjects ($M = 69{\cdot}42$). Pictures were recognized *less* readily ($M = 67{\cdot}96$) than concrete words ($M = 71{\cdot}58$) or abstract words ($M = 70{\cdot}67$). The significant interaction revealed the superiority of high-imagers in all stimulus attribute conditions, the greatest effect occurring with the picture list where differences reached significance at the 0·01 level ($t = 3{\cdot}69$; d.f. = 22). Differences between imagers were not significant with either the concrete or abstract word lists.

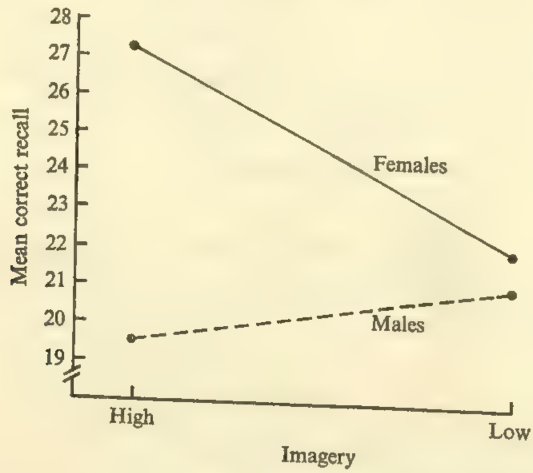


Fig. 2. Incidental recall as a function of imagery ability and sex.

Incidental learning. The significant main effects ($P < 0{\cdot}001$) obtained by analysis of variance were sex ($F = 14{\cdot}73$; d.f. = 1, 60) and stimulus attribute ($F = 109{\cdot}94$; d.f. = 2, 60). Incidental learning was significantly higher for females ($M = 24{\cdot}61$) than males ($M = 20{\cdot}33$) and for pictures ($M = 34{\cdot}04$) relative to concrete words ($M = 18{\cdot}12$) and abstract words ($M = 15{\cdot}25$). Although imagery did not emerge as a significant main effect ($F = 2{\cdot}87$; d.f. = 1, 60), it did interact with sex ($F = 9{\cdot}55$; d.f. = 1, 60; $P < 0{\cdot}01$). Plotted in Fig. 2, this interaction is almost identical to the one previously obtained in a PA task (Ernest, 1968). In the present study, high-imagery females obtained significantly higher incidental learning scores than high-imagery males ($t = 4{\cdot}76$; d.f. = 34; $P < 0{\cdot}001$). Of even greater interest, however, is

the finding that the difference between high- and low-imagery ability was significant in the female sample ($t = 3{\cdot}29$; d.f. $= 34$, $P < 0{\cdot}01$) but not the male sample ($t = 0{\cdot}96$; d.f. $= 34$).

Discussion

The relationship between imagery ability and incidental learning was confirmed with female subjects in Expt. II. The interaction between imagery level and sex suggests that, in some tasks, females 'use' imaginal processes to facilitate recall whereas males do not. This interpretation is supported by the results obtained with an Individual Differences Questionnaire* which measures the degree to which an individual typically employs verbal and or imaginal techniques in thinking, studying and problem-solving. High-imagery females obtained significantly higher imaginal scores than low-imagery females ($t = 3{\cdot}44$; d.f. $= 34$; $P < 0{\cdot}01$). Differences were not significant with males, nor in any comparisons involving the verbal scores. Imagery ability and incidental learning were not related in Expt. I, possibly due to the difficulty of the task.

Learning differences between high- and low-imagers were not restricted to females, however. In a free-recall task (Expt. I) high-imagery males were superior to low-imagery males in recalling words. The results for females were in the reverse direction (see Fig. 1), which is consistent with Stewart's (1965) findings with a sample consisting solely of females. She found that low-imagers exceeded high-imagers in the free recall of verbal materials and accordingly interpreted the behaviour of low-imagers in terms of verbal (as opposed to imaginal) skills. Such an intepretation is questionable when viewed in the light of present findings involving both sexes.

The effects of stimulus attribute in the recall task of Expts. I and II are consistent with Paivio's (1969) two-process theory of memory. It is particularly noteworthy that, although pictures were recognized less readily than concrete or abstract words in Expt. II (a result probably due to the transformation required in providing a verbal label for the pictures), they surpassed words in the subsequent incidental recall task. Whether the superior incidental recall of the colour component of the pictorial items relative to the verbal items in Expt. I can be atttribued to the nature of the stimuli *per se* or to the fact that colour is normally more intrinsic to pictures than to words cannot be answered by this study.

Interestingly, imagery ability emerged as a significant factor in the recognition and labelling of pictorial stimuli. This finding suggests that memory representations of the stimuli, established through previous perceptual experiences, and/or their verbal labels, were more 'available' to high-imagers than low-imagers, thus providing further evidence for the functional significance of imagery in tasks involving a memory component.

This research was supported by a National Research Council (NRC) of Canada Postgraduate Scholarship to the first author, an NRC Grant (APA-87) to the second author, and the University of Western Ontario Research Fund.

* This questionnaire, devised within our laboratory, has not as yet been published.

REFERENCES

CHRISTIANSEN, T. (1969). Sex and level of mediator abstractness as variables in mediate association. *Psychon. Sci.* **14**, 160, 162.

ERNEST, C. H. (1968). Individual differences in imagery ability and compound stimulus effects in paired-associates learning. (Unpublished master's thesis, University of Western Ontario.)

ERNEST, C. H. & PAIVIO, A. (1969). Imagery ability in paired-associate and incidental learning. *Psychon. Sci.* **15**, 181–182.

PAIVIO, A. (1969). Mental imagery in associative learning and memory. *Psychol. Rev.* **76**, 241–263.

PAIVIO, A., YUILLE, J. C. & MADIGAN, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *J. exp. Psychol.* (Mongr. Suppl.) **76**, 1–25.

SHEEHAN, P. W. & NEISSER, U. (1969). Some variables affecting the vividness of imagery in recall. *Br. J. Psychol.* **60**, 71–80.

STEWART, J. C. (1965). An experimental investigation of imagery. (Unpublished doctoral dissertation, University of Toronto.)

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York: Bureau of Publications, Teachers' College.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

(*Manuscript received* 1 *January* 1970)

Br. J. Psychol. (1971), 62, 1, pp. 73–79
Printed in Great Britain

THE REPETITION OF SINGLE SPEECH SOUNDS UNDER DELAYED AUDITORY FEEDBACK

BY JANET F. FLETCHER AND AUBREY J. YATES
Department of Psychology, University of Western Australia, Nedlands, Western Australia

Three experiments were carried out to determine the effect of delayed auditory feedback (DAF) on the rate of repetition of single speech sounds. Other variables investigated were variation in repetition rate among speech sounds under DAF, and the effect of the subject's sex on repetition rate. Subjects were found to repeat sounds significantly fewer times under DAF. There was no significant variation among sounds repeated under DAF, nor was there a sex effect. Since the major finding was in the opposite direction to that reported by Chase (1958), a fourth study was carried out to investigate a possible cause of this discrepancy.

It has generally been assumed that speech is monitored at its most basic level on the basis of a 'unit' of speech (Lee, 1951; Fairbanks, 1954; Chase, 1958; Black, 1955, 1958). Hypotheses concerning the nature of such a unit have ranged from the suggestion that the unit is a 'semi-periodic, relatively long articulatory cycle, with a correlated cycle of output' (Fairbanks, 1954) to specific hypotheses that the sound and the syllable are crucial units in the monitoring of speech (Black, 1955). Little empirical evidence, however, has been put forward in support of the various hypotheses.

Chase (1958) preferred to make no assumptions concerning the nature of the unit involved in speech monitoring, but he suggested that when a person speaks under conditions of delayed auditory feedback (DAF), the changes that occur in subsequent speech (Yates, 1963) result from the facilitation of the circulation and recirculation of the speech unit in the speech–hearing feedback loop. To investigate this hypothesis, Chase required 20 subjects to repeat the speech sound [b] as quickly as possible for 5 sec. under conditions of synchronous auditory feedback (SAF) and DAF (216 msec. delay). The intensity of feedback in both conditions represented a gain of 20 db to the subject's speech. Chase reported that 75 per cent of his subjects did repeat the sound more quickly under DAF. These findings have been interpreted as evidence that DAF facilitates the circulation and recirculation of the speech unit. All that can properly be said, however, on the basis of these results, is that DAF facilitated the circulation and recirculation of one particular speech unit, the phoneme [b]. It is not known whether these findings would hold in the case of vowels or other consonants of a different phonetic nature. The present series of experiments attempted to replicate and extend Chase's study.

EXPERIMENT I

The aims of this experiment were to determine whether the rate of repetition of speech units as small as single speech sounds is significantly affected by DAF; whether there is any significant variation in repetition rate among sounds under DAF; and whether there is any significant difference in the rate of repetition of single speech sounds under DAF as a function of sex of the subject.

Method

Subjects. The subjects were 50 students (25 males, 25 females) who volunteered for the experiment and who were obtained from students enrolled in 1967 at the University of New England and Armidale Teachers' College. No subject had experienced DAF before and all were screened for speech and/or hearing disorders or irregularities.

Apparatus. The subject was seated at a table and spoke into a Shure Model 545 Unidyne III dynamic microphone. The microphone was on a stand such that it was at mouth level of the average subject when seated. A small red light was visible to the subject. From the microphone the verbal output was transmitted to an Akai M8 magnetic tape-recorder, where it was continuously recorded, and to a locally constructed DAF unit from which it was transmitted, either delayed or not delayed, to the subject through Telephonic TDH-39 earphones set in an Auraldome Model 7AR sound-isolating headphone. The intensity of the feedback in both DAF and SAF conditions was maintained at a gain of 35 db above the subject's vocal output by means of the amplifier gain control of the unit.

Table 1. *Mean repetition rate of consonants per 5 sec. period summed over subjects under SAF and DAF in Expt. I*

Sound	SAF	DAF
ch as in choke	15·64	14·40
su as in measure	12·44	11·48
n as in nail	12·32	11·88
m as in mail	12·18	11·16
th as in then	12·30	10·92
sh as in show	14·06	12·76
t as in tale	15·84	15·46
h as in help	14·78	12·54
f as in fine	13·96	12·80
th as in thin	13·12	12·24
g as in game	15·02	14·76
v as in vine	11·82	11·76
j as in joke	15·62	13·98
c as in came	15·78	14·70
l as in long	12·36	11·74
d as in dog	15·86	15·26
b as in bet	15·46	15·06
s as in seal	13·92	12·58
p as in pet	15·50	14·66
z as in zeal	12·32	12·16
Average of means	14·01	13·12

Experimental design and procedure. The subjects were tested individually, and the complete testing of each subject was carried out in a single test session. The subject was required to repeat 20 different consonants (Table 1) as quickly as possible for 5 sec. each, under both SAF and DAF (200 msec. delay), a total of 40 repetitions. Random combinations of sounds and delays were generated for the purpose. A different random order of the 40 combinations was used for each subject.

Before the test repetition of each sound, the subject was given a demonstration of the sound by the experimenter and was required to repeat it until the experimenter was satisfied that the sound was being said correctly. Before this practice trial and also before the test run, the subject was given a rest of 15 sec. to minimize fatigue.

The following instructions were given to the subject: 'This is an experiment in human communication. During the course of the experiment you will be asked to speak into the microphone while wearing headphones. The speech task that you will be required to perform is to repeat a single sound as instructed as quickly as possible for 5 sec. Before each trial you will be shown the sound to be repeated and asked to repeat it after me. When you have repeated the sound correctly, your earphones will be put on, and you will have 15 sec. rest. When the red light

on the table in front of you lights up, you are to repeat the sound as quickly as possible until the red light comes on a second time. Remember to repeat the sound exactly as you practised it. Although speed is important it is also vital that you repeat the sound correctly.'

The subject was then given a card on which was printed the first sound, and the sound was demonstrated by the experimenter. The experimenter said the sound only once at [illegible] so as not to provide a set with regard to speed of repetition. When the sound had been correctly executed, the subject was fitted with headphones. Testing began at the onset of the red desk light, 15 sec. after the completion of the last practice trial as indicated by a stopwatch. At the conclusion of testing, the subject was requested not to discuss the experiment with anyone.

Analysis of performance. The number of sound repetitions per 5 sec. period was assessed by recording the repetitions at a speed of 7½ in. per second and playing them back at 3¾ in. per second so that the actual number of sounds could be counted by the experimenter. (To check the reliability of the scoring technique the complete records of 10 per cent of the subjects, chosen at random, were rescored several months after the completion of the original scoring. The correlation between the two sets of scores was +0·992.)

The difference between number of repetitions under SAF and DAF (SAF − DAF) for each sound for each subject was then calculated. A positive difference score indicated that the subject had slowed down under DAF in relation to performance under SAF, whereas a negative difference score indicated that the subject had speeded up under DAF.

The difference scores obtained were subjected to an analysis of variance. The analysis was a minor variation on the standard 2 × 2 split plot design, with sex in whole plots and sounds in subplots.

Results

The analysis of variance (Table 2) indicates a significant effect of delay resulting from the subjects repeating sounds fewer times under DAF. However, there was no significant differential effect of DAF on sounds, nor was any significant effect of sex found. The sound by sex interaction was not significant.

Table 2. *Analysis of variance of difference scores. Expt. I*

Source	D.F.	S.S.	M.S.	*F*	*P*
Delay	1	810·00	810·00	80·677	< 0·01
Sex	1	11·24	11·24	0·196	n.s.
Whole-plot error	48	2751·66	57·33	—	—
Sound	19	286·40	15·07	1·50	n.s.
Sound × sex	19	207·06	10·90	1·085	n.s.
Subplot error	912	9157·62	10·04	—	—

Discussion

The finding that DAF produced a highly significant change in the number of sounds repeated in a 5 sec. period was expected. The difference, however, was in the opposite direction to that reported by Chase (1958) in his experiment involving the repetition of the single sound [b]. Furthermore the mean difference scores for each sound summed over subjects showed that this effect was the same for all sounds, including the sound [b].

It was noted, however, that although subjects repeated the sounds correctly during pre-experimental practice, performance under both SAF and DAF conditions was considerably less accurate, as judged by the experimenter. The essential phonetic nature of the sounds was frequently altered, and vowel nuclei were sometimes inserted between repetitions of the consonants. This effect could have made the difference

score (SAF – DAF) less meaningful and lessened the chances of obtaining significant between-sounds variation.

In Expts. II and III an attempt was made to overcome these difficulties.

EXPERIMENTS II AND III

In these experiments several methodological changes were introduced. The subject was required to practise repeating the sound as quickly as possible before the test repetition of each sound in an effort to minimize the variation of each sound. Since no significant between-sound variations were found in Expt. I the number of consonants in Expt. II was reduced from 20 to 12 in order to reduce the effects of fatigue and boredom by shortening the length of the testing session. Expts. II and III attempted to replicate the findings of Expt. I and also to determine whether there is any significant variation in rate of repetition under DAF between consonants and vowels.

Method

Subjects. One hundred subjects were obtained from students enrolled at the University of New England in 1968. Fifty subjects (25 male, 25 female) participated in each experiment. All subjects were naive with respect to DAF and were screened for speech and/or hearing disorders or irregularities.

Apparatus, experimental design and procedure. Apparatus, experimental design and procedure

Table 3. *Mean repetition rate per 5 sec. period summed over subjects of speech sounds in Expts. II and III*

Sound		SAF	DAF
	Experiment II: Consonants		
g	as in game	15·95	15·30
v	as in vine	11·65	11·35
c	as in came	15·65	15·70
th	as in thin	13·35	12·95
j	as in joke	15·60	15·20
n	as in nail	12·65	12·20
sh	as in show	14·30	13·45
b	as in bet	16·35	16·35
f	as in fine	13·90	13·35
th	as in then	11·90	11·15
m	as in mail	12·00	11·95
p	as in pet	16·15	16·65
Average of means		14·12	13·80
	Experiment III: Vowels		
a	as in father	17·85	17·15
ea	as in seat	18·10	17·65
i	as in sit	19·45	19·45
ur	as in fur	17·60	16·40
o	as in lot	18·85	19·10
oo	as in look	18·65	18·50
a	as in ago	18·05	18·40
ew	as in crew	18·20	17·25
u	as in hunt	18·80	19·15
aw	as in law	17·40	16·90
e	as in set	19·05	18·55
a	as in sat	19·00	13·33
Average of means		18·42	18·07

were the same as in Expt. I with the exception of the methodological changes noted above. The consonants and vowels repeated by the subjects in Expts. II and III respectively are listed in Table 3.

Analyses of performance. The number of sound repetitions per 5 sec. period was assessed in the same manner as in Expt. I. (Rescoring a random 10 per cent of the data in both studies after an interval of several months yielded a reliability coefficient of 0·997 in the case of Expt. II and 0·994 in the case of Expt. III.)

The difference between number of repetitions under SAF and DAF (SAF − DAF) for each sound for each subject was calculated and the difference scores obtained in Expts. II and III subjected to a split plot analysis of variance. A further analysis of variance was carried out to compare the difference scores obtained with consonants with the difference scores obtained with vowels.

Results

The analyses of variance indicate that DAF was accompanied by a significant ($P < 0{\cdot}025$) change in the repetition rate of a speech unit as small as a single speech sound, due to the subjects repeating sounds fewer times under DAF. This effect, however, was not as pronounced as in Expt. I, and when the data were analysed separately for consonants and vowels the results were not significant. There were no differential effects due to sex, either alone or in interaction. No differential effects in number of sounds repeated under DAF within consonants or within vowels were found; nor were there any significant differences between number of consonants and number of vowels repeated under DAF.

Discussion

The findings of Expts. II and III were very similar to those of Expt. I, despite alterations in the experimental method to ensure greater accuracy in pronunciation of sounds.

Variation in repetition rate under DAF was again produced by slower repetition rate under DAF, throwing serious doubt on Chase's (1958) hypothesis that the effects of DAF are due to facilitation of the circulation and recirculation of speech units in the speech–auditory feedback loop.

One possible reason for the difference between the results of the present studies and those obtained by Chase (1958) may lie in the instructions given to the subject. Chase instructed his subjects to 'repeat the sound as quickly as possible', whereas in the present studies both speed and accuracy were stressed. An examination of scores in the present studies suggested that some subjects may have considered either speed or accuracy to be more important. However, even subjects who seemed to stress speed never equalled the rate maintained by the majority of Chase's subjects.

It is possible, therefore, that a negative relationship exists between rate of repetition under SAF conditions and extent of breakdown under DAF, in which case Chase's subjects (who on the sound [b] averaged 31·7 repetitions per 5 sec. period under SAF) may have come from one end of the SAF continuum, whereas subjects in Expts. II and III (who averaged over all sounds 16·27 repetitions per 5 sec. period) may have come from the other end of the SAF continuum.

Mean SAF scores and difference scores (breakdown measure) on Expts. II and III combined were correlated to examine this hypothesis. The correlation ($r = 0{\cdot}31$; $P < 0{\cdot}005$, one-tailed test) indicated the need to examine the hypothesis that 'fast' subjects under SAF speak faster under DAF, and 'slow' subjects speak more slowly.

EXPERIMENT IV

Since in Expts. II and III the mean repetition rate of even the fastest subjects never equalled the rate maintained by the majority of Chase's subjects, it was impossible to achieve the above aim by using groups of subjects who varied in their 'natural fastest rate' under SAF and then comparing their performances under DAF. Instead, two groups were trained to repeat the sounds at the desired rate. This was done by pacing the subjects in a practice period by the amplified clicks of an oscillator set at 6 Hz (fast group) or 3 Hz (slow group), since six sounds per second was the mean rate of Chase's (1958) subjects, and three sounds per second was the mean rate of the subjects in Expts. II and III.

Method

Subjects. Forty-eight subjects (24 males, 24 females) were obtained from students enrolled at the University of New England in 1968. They were divided into two equal groups (12 males, 12 females in each), one of which became the fast experimental group, and the other the slow experimental group. All subjects were naive with respect to DAF and were screened for speech and/or hearing disorders or irregularities.

Table 4. *Mean repetition rate per 5 sec. period summed over subjects and sounds for fast and slow conditions*

	Fast group		Slow group	
	SAF	DAF	SAF	DAF
Mean summed over sounds	339·83	337·17	187·96	193·58
Standard deviation	25·27	21·46	18·28	28·77
t (d.f. = 23)	0·887 (n.s.)		1·595 (n.s.)	

Apparatus. The apparatus was identical to that outlined in Expt. I with the addition of a Kikusui Denpa Co. Ltd. Model 415 K-C Decade oscillator, to which was attached an amplifier.

Experimental design and procedure. The experimental design and procedure were basically the same as those outlined in the description of Expt. I. Several modifications were, however, introduced. Plosives (p, t, b, c, d, g) were the only sounds repeated by the subjects in this experiment. The limited range was justified in view of the results of Expts. II and III where no significant between-sound variations were observed.

Before the test repetition of each sound, the subject was given a demonstration of the sound by the experimenter, and the repetition rate was indicated by means of the oscillator. The subject was then given 9 sec. in which to practise the sound, the first 3 sec. being actual pacing by the oscillator, the next 3 sec. repetition without the oscillator, and the final 3 sec. again pacing with the oscillator. By means of this procedure, subjects learned to maintain the required rate quite accurately. Subjects were instructed to repeat the sound 'exactly as practised' during test trials.

Analysis of performance. The number of sound repetitions per 5 sec. period was assessed in the same manner as in the previous experiments. (A repeat scoring of a random 10 per cent of the data after an interval of several months yielded a reliability coefficient of 0·997.) SAF and DAF scores thus obtained were summed over sounds for each subject, to give a total of 24 SAF and DAF scores for both fast and slow experimental groups. A t test was then carried out on the scores from both groups.

Results

The t tests (Table 4) indicate that there was no significant difference between SAF and DAF scores in either fast or slow groups. Hence the fast subjects under SAF did not become faster under DAF, nor the slow subjects under SAF slower under DAF.

Discussion

It would appear from the above results that there is no evidence for the hypothesized relationship between performance under SAF and speech breakdown under DAF—provided that Expt. IV does actually provide an adequate test for the hypothesis. It is possible that there may be a crucial difference between subjects whose 'natural fastest rate' is three or six sounds per second and those who are taught this rate. In this case Expt. IV may have introduced a confounding variable. This would seem to be a likely explanation, since in Expts. II and III where rate was natural fastest, there was a significant, if low, correlation between rate under SAF and breakdown scores under DAF, whereas in Expt. IV, with subjects drawn from the same population as in Expts. II and III, but endeavouring to repeat sounds at a set rate, *t* tests showed no evidence of a relationship between performance under SAF and DAF.

The absence of a significant difference between performance under SAF and performance under DAF parallels the results found by Chase *et al.* (1961) when they had subjects performing a regularity task in keytapping. Chase concluded from this study that a certain amount of temporal complexity in a task was necessary before breakdown in performance under DAF occurs. It is possible that the practice trials in Expt. IV, when subjects were paced by the amplified oscillator clicks, functioned to lessen the temporal complexity of the task to a level where breakdown in performance ceased to occur.

It would thus be worthwhile to re-examine the hypothesis of Expt. IV, if a population could be found ranging in their 'natural fastest rate' from three to six sounds per second.

This research was facilitated by a grant from the Australian Research Grants Committee. The studies were carried out at the University of New England, New South Wales, Australia.

References

Black, J. W. (1955). Studies of delayed sidetone. *Pubblicazioni Dell' Universita Cattolica Del Sacro Cruore*, n.s. **44**, 3–23.

Black, J. W. (1958). The reading of messages of different types and number of syllables under conditions of delayed sidetone. (Joint Res. Proj. NM 180299 Subtask I Report 78 U.S. Naval School of Aviation Medicine.)

Chase, R. A. (1958). Effect of delayed auditory feedback on the repetition of speech sounds. *J. Speech Hear. Disorders* **23**, 583–590.

Chase, R. A. *et al.* (1961). Studies on sensory feedback. II. Sensory feedback influences on keytapping motor tasks. *Q. J. exp. Psychol.* **13**, 153–167.

Fairbanks, G. (1954). Systematic research into experimental phonetics. I. A theory of the speech mechanism as a servosystem. *J. Speech Hear. Disorders* **19**, 133–139.

Lee, B. S. (1951). Artificial stutter. *J. Speech Hear. Disorders* **16**, 53–55.

Yates, A. J. (1963). Delayed auditory feedback. *Psychol. Bull.* **60**, 213–232.

(*Manuscript received* 6 *December* 1969)

Br. J. Psychol. (1971), **62**, 1, pp. 81–88
Printed in Great Britain

QUALITATIVE GUSTATORY CHARACTERISTICS OF DISODIUM-5′-GUANYLATE

By R. A. M. GREGSON and M. B. SIMMONDS

Department of Psychology and Sociology, University of Canterbury, Christchurch, New Zealand

Descriptions of the taste qualities of a nucleotide, disodium-5′-guanylate, are reported under two different response sets, lenient and stringent. Control comparison stimuli were de-ionized water and weak sodium chloride, and near threshold taste concentrations were used. Twenty subjects, 10 under each response set, participated. D5′G has a complex taste, and induces a marked false positive detection response pattern which resembles its own taste. The implications for methodology in gustatory psychophysics are noted.

The molecular structures of RNA and DNA have become familiar even to workers outside biochemistry because of their importance in genetics. In particular, psychologists concerned with the possibility of the inheritance of learnt behaviour patterns and the microphysiology of information storage in the central nervous system have speculated at length on the relevance of their structures for encoding learnt responses. Less well known is the role of RNA in furnishing the building blocks of a series of taste stimuli whose precise psychophysical qualities have proved singularly difficult to identify.

The nucleic acids are joined in a chain formation; each component is a nucleotide made up of a base, plus a sugar (ribose), plus a phosphoric acid link. Some nucleotides can exist alone naturally and not as links in RNA. If the chain is broken at a phosphordiester link, either by the action of enzymes produced by micro-organisms (Kuminaka *et al.*, 1964) or by some industrial process, a number of isomers of the links can occur, but only the 5′-ribonucleotides have taste properties, especially those having an OH group in the 6 position of the purine nucleus (Eberhardt, 1964). The sodium salts disodium-5′-inosinate and disodium-5′-guanylate have a characteristic 'meat-like' taste which they share with the more widely known taste additive monosodium glutamate (MSG). MSG was the earliest widely available substance with this taste and has been studied for a number of diverse psychophysical and behavioural effects which it apparently induces. The nucleotides ingested in large quantities may have side-effects which are not typical of more usual flavouring agents. Penau (1966) asserts that 5′-inosinic acid and 5′-guanylic acid are non-toxic, but a potentially lethal condition associated with MSG, and originally named Kwok's syndrome (Kwok, 1968; Ambos *et al.*, 1968) has been identified; a minority of human subjects are reported to suffer incapacitating autonomic effects after ingesting MSG at the concentrations commonly used in normal food seasoning.

Attention has been centred on the action of MSG and D5′I as gustatory threshold depressers (Konusu *et al.*, 1960). It has been argued that if these substances enhance taste intensity they do so by effecting a temporary increase in sensory acuity as well as by having a taste of their own. It seems generally agreed that MSG and the nucleotides have no smell, so they are suitable, if pure, for taste research. Early studies used MSG contaminated by one or more amino acids which were not separated

during manufacture, but the salts of nucleotides now available are relatively pure and stable.

The way in which MSG and the nucleotides function as flavour enhancers is not really understood; they have a taste of their own (Yoshida *et al.*, 1966; Yoshida & Saito, 1969); they are also asserted to modify other tastes, but this evidence is contradictory (Crocker & Henderson, 1932; Tseng & Chu, 1933; Crocker & Sjöström, 1948; Lockhart & Gainer, 1950; Mosel & Kantrowitz, 1952; Van Coff *et al.*, 1954; Pilgrim *et al.*, 1955).

The problem here is that the nucleotides have a taste which is, at near-threshold intensities, reducible to a mixture of the four primary taste sensations, and in a multidimensional scaling in Euclidean space by Yoshida (1963) is apparently located at a unique vector in three dimensions. But as a complication the nucleotides also have a capacity at some concentrations to elicit sensations described as 'mouth filling' (or in Japanese as 'umami', translated as 'delicious') which are not strictly tastes in either phenomenological or physiological terms. In some of the cited experiments the nucleotide has been introduced at 'subthreshold' intensities, in combination with other taste stimuli, but it is not in fact clear what this means, when the psychophysical procedures used do not identify the patterns of true and false positive detection responses to be expected under all the stimulus presentations employed. When a substance like D5′G gives rise to complex taste sensations, in a pattern which may change with the stimulus concentration, it is of trivial value to be told that it is subthreshold at an undefined response criterion on a unidimensional psychophysical continuum.

The psychophysical effects of the nucleotides are supposedly optimal when they are at very weak concentrations (Shimazoho, 1964). The upper psychophysical threshold is reached soon after the detection concentration, and our own observations suggest that there may even be a turning point in the concentration–sensation intensity function with increased concentration. The optimal action of D5′I and D5′G is claimed (Adachi *et al.*, 1965; Yamaguchi, 1965; Sato *et al.*, 1966) to be synergistic upon MSG: this leads to the use of a mixture of 1:20 of D5′I:MSG. However, the very high cost of D5′I and D5′G is partly the reason for this, although D5′I has a lower threshold than MSG, and D5′G is possibly even lower. Amerine *et al.* (1965) give the range of threshold concentrations of the 5′-nucleotides as 0·0035–0·025 per cent, and D5′G in this study is at the lower end of the range.

The studies mentioned are inconclusive both with reference to any possible gustatory detection threshold shift induced by the nucleotides, and to the actual qualitative taste character elicited by the nucleotides, above or near threshold. The inconclusiveness stems from the diversity of psychophysical methods used, and the incomplete data presentation (Gregson, 1965). It has therefore been thought necessary, in examining D5′G as the least documented and potentially most active nucleotide in a psychophysical sense, to revert to a simple and basic exploratory study. This was conducted to obtain a broad qualitative description of the taste and after-taste of D5′G, and an estimate of its threshold concentration region, and to control for the qualitative pattern of false positive responses. None of the previous studies we have reviewed has done this in sufficient detail.

METHOD

Procedure

The experiment was performed in the University of Canterbury Gustatory Psychophysics Laboratory. See Gregson (1964) for relevant details of the construction of this unit.

Seven taste stimuli were prepared: pure water, 0·12% NaCl, 0·27% NaCl, 0·005% D5'G, 0·010% D5'G, 0·020% D5'G, and 0·040% D5'G.

The D5'G was supplied in the powder form by the Kyowa Hakko Kogyo Co. Ltd. of Tokyo, Japan, to whom our thanks are due.

The water used, both in the preparation of stimulus solutions and for rinsing of the subject's mouth between each actual, was deionized to a resistance greater than 1·5 megohm per cm which is equivalent to triple distillation. The purity of the source water has also been checked; it is relatively free from organic and inorganic contamination (Gregson, 1964).

Table 1. *Order of presentation of the 36 stimuli*

Trial	Series 1 and 2	Series 3 and 4	Series 5 and 6
1	Water	Water	0·005 % D5'G
2	Water	Water	0·010 % D5'G
3	Water	0·005 % D5'G	0·020 % D5'G
4	Water	0·010 % D5'G	0·040 % D5'G
5	0·12 % NaCl	0·020 % D5'G	Water
6	0·27 % NaCl	0·40 % D5'G	Water
	Rinse	Rinse	Rinse

A strict time sequence for responding was signalled to the subjects by coloured lights, and was as follows: 1–4 sec. (yellow signal light) – prepare to taste; 5–15 sec. (red signal light) – taste and make responses for 10 sec., by key pressing whilst stimuli are held in the mouth; 16–60 sec. (green signal light) – spit out solution in cuspidor, rinse with water (if instructed to do so) and rest before presentation of the next stimulus.

Solutions were presented to the subjects in standard 50 ml. beakers, each containing 10 ml. of solution. All samples were stored in a water bath at 25° C for not less than 15 min. before presentation to the subject. Responses were made on a bank of Morse keys, which were connected via Massey–Dickinson modules to an Esterline Angus recorder. Each key corresponded to a single response statement. The response statements simultaneously available to the subject were: no taste discernible, a very faint unidentifiable taste, a faint unidentifiable taste, a very faint sweet taste, a faint sweet taste, a very faint salty taste, a faint salty taste, a very faint sour (acid) taste, a faint sour (acid) taste, a very faint bitter taste, and a faint bitter taste.

Solutions were presented in racks containing three beakers, and at no time were the subjects informed by the experimenter whether their responses were 'correct' or not. (The *a priori* definition of 'correct' raises considerable difficulties in an experiment of this kind.) Each subject took about 40–45 min. to complete the experiment. The order of presentation of stimuli, 36 in all, was the same for all subjects, and is shown in Table 1.

The low concentrations of NaCl in series 1 and 2 were included to avoid the appearance, to the subjects, of a completely null stimulus series on the early trials of the experiment, and to provide a control, as the taste of NaCl has been thoroughly studied. The water blanks on the first two trials of series 3 and 4, and on the last two trials of series 5 and 6 provide a means of comparison between false positive responses at the beginning and end of an ascending series of D5'G, thus providing a partial measure of any after-taste effects which D5'G may possess. We thus have water presented in three different contexts: (1) Water preceding NaCl (column 1 of the relative frequency matrices represents pooled data from trials 1–4 of series 1; likewise column 2 of the graphs represents trials 1–4 of series 2). (2) Water preceding D5'G (columns 3 and 4 of the graphs represent trials 1 and 2 pooled from both series 3 and series 4 respectively). (3) Water following D5'G (columns 5 and 6 represent trials 5 and 6 pooled for both series 5 and series 6 respectively).

Instructions

Two different sets of instructions to the subjects were presented along with details of a general procedure for responding.

Lenient instructions (*set A*). 'Read these instructions carefully in full. In this experiment you will be presented with small samples of liquids in beakers. The liquids are mostly water, but with small traces of various taste substances added. It should be possible, in many if not in all samples, to detect some very weak tastes. These weak tastes may only appear for a moment, and then fade away. They may only seem to appear in one part of your mouth and not in another. They may come and go as you move your tongue around. You will be required to hold each sample in your mouth, without swallowing, for about 10 sec. During this time you must press a key as appropriate every time you think you experience a taste sensation, however slight or short-lived it is. You may get a variety of sensations from one sample during the time it is in your mouth. Be prepared for this to happen and make as many different or repeated key presses as are necessary to give a full record of your sensations. It is better to guess than not to make an answer at all.'

Stringent instructions (*set B*). 'Read these instructions carefully in full. In this experiment you will be presented with small samples of liquids in beakers. The liquids are mostly water. It should be possible in some but not all of the samples to detect very weak tastes. You will be required to hold each sample in your mouth, without swallowing, for 10 sec. Each time you are *certain* you experience a taste sensation, which persists for more than a moment, you should press the key which comes nearest to describing your sensation. You do not have to press a key for a sample if you do not experience any tastes. You may make more than one response to a sample while it is in your mouth. *Please do not guess.*'

Subjects

The subjects were 20 first- and second-year undergraduate psychology students, naive with respect to the psychophysical theory involved, and serving to fulfil their laboratory course requirements. Ten subjects were used under each of the two sets of instructions, and no subject served under both conditions. There are thus two separate and parallel experiments, enabling comparisons to be made between the data collected under two different response sets.

RESULTS

The responses have been treated as samples from a multinomial response generating process, in which the basic response categories are six, formed by pooling the response categories 'very faint' and 'faint' for each qualitative response available to the subjects. This gave the following categories: no taste discernible, unidentifiable taste, sweet, salty, sour (acid), and bitter.

The responses have been pooled over subjects, but the stimulus types only partially pooled; the six runs where water blanks occur in the total stimulus series have been analysed separately, as indicated in the section on procedure. They thus furnish six control blocks on false positive response rates to water during the progress of the experiment.

The two relative frequency matrices, for instructions A and B respectively, are shown in Figs. 1 and 2.

The comparison of the response profiles for the various stimuli distinguished in Figs. 1 and 2 cannot readily be effected by available statistical models without making some prior assumptions about the mode of response generation which are certainly false. An approximate method, following Good (1965), has been developed by the senior author, but will be reported elsewhere because of its specialized nature.

It is seen that the form of instruction makes a considerable difference to the

qualitative psychophysics of the situation. If the stringent instructions B are used, the data are sufficiently orderly for the profiles on D5′G to show the linear trend on true detection probability with an increase in log stimulus concentration, as is commonly found in gustation over the near-threshold range. At the maximum

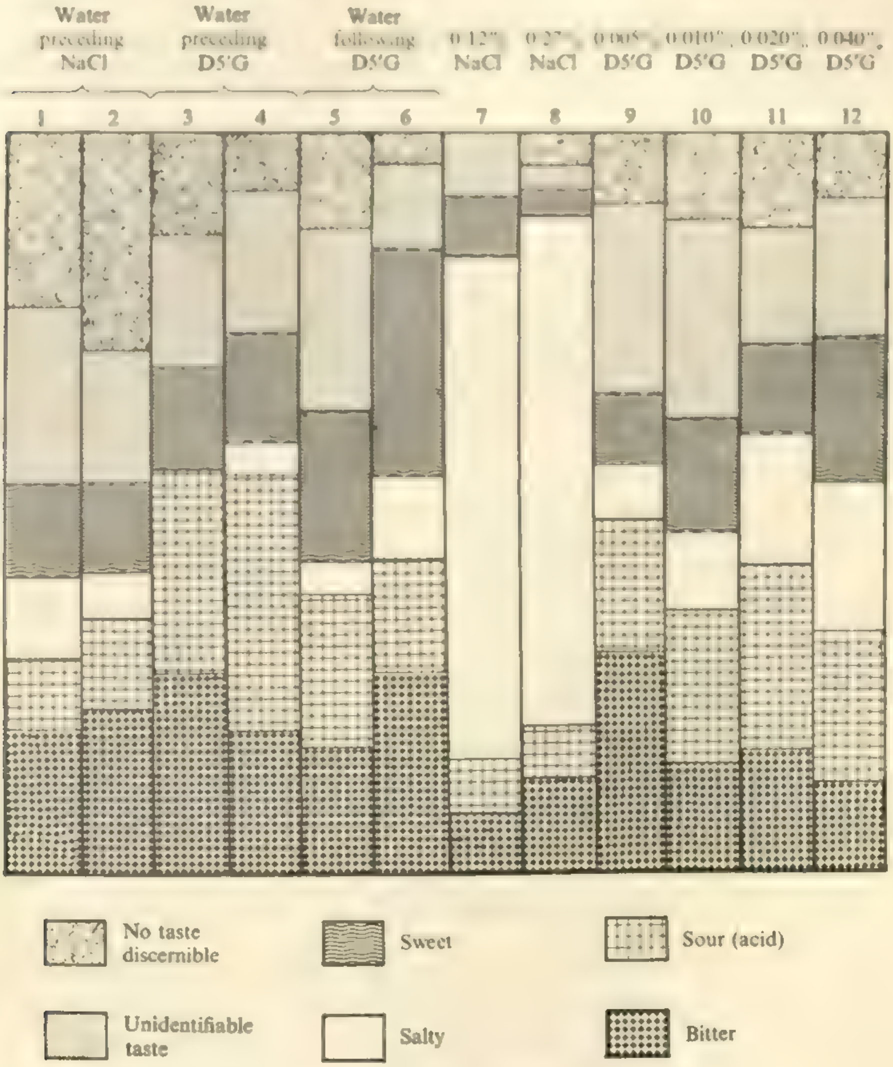


Fig. 1. Relative frequency matrix for lenient instructions (set A).

concentration, 0·04%, D5′G is clearly qualitatively different from the weakest D5′G considered, and from the very clear NaCl profiles, which show the sweet taste of weak NaCl reported by van Skramlik (1926) at the near-threshold concentrations we deliberately chose here. The false positive responses for water were, however, so dominated by the persistent after-taste of D5′G in the later stages of the experiment that the signal–noise separation of D5′G to water became borderline.

D5′G has a complex taste, as Yoshida & Saito (1969) report for the amino acids, but it has components of all four primary tastes and does not exclude salty as they suggest. It is, of course, very different from the stronger NaCl but has some slight similarity with the weaker NaCl profile.

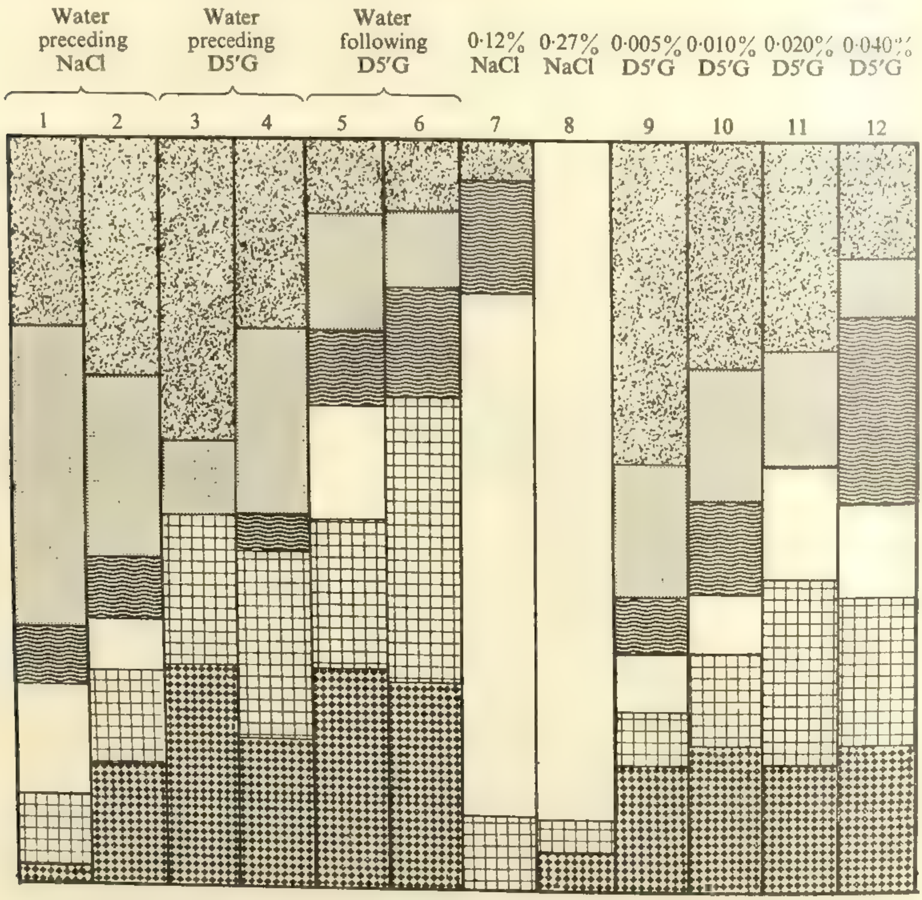


Fig. 2. Relative frequency matrix for stringent instructions (set B). Key as in Fig. 1.

There is an almost complete absence of 'salty' responses in columns 3 and 4 of both Figs. 1 and 2. The prior presentation of NaCl shifts both the adaptation level for detecting 'salty' and the false positive detection rate, though these two processes are confounded to some extent.

Part of the increased noise in condition A (lenient set) may be due to the method of pooling data. In condition B most subjects gave only one response to each stimulus, with an average of 1·21 responses per stimulus, whereas in condition A there was an average response rate of 1·55. If only the first responses to stimuli were analysed the differences between Figs. 1 and 2 would be less marked, because subjects tend not to repeat the same response description when they make multiple responses. Whether this represents an information loss of real positive or false positive detection responses has not, so far as we know, yet been explored in gustatory psychophysics.

The big difference in the overall pattern, for the two instructions A and B, emphasizes the difficulties which arise when this source of variation is not controlled. There is so much potential variability in the qualitative perception of both weak tastes and after-tastes that comparison between studies which do not control for this feature is rendered impossible. It is not just that the subject changes his readiness to report a taste, as a simple signal-detection theory paradigm would suggest, but his relative use of different qualitative descriptions changes, and his multivariate noise distribution changes in qualitative character and modal intensity with time. This pattern is consistent with similar results on sodium benzoate (Gregson, 1969).

Conclusion

Disodium-5′-guanylate has a complex taste and a complex after-taste which is similar, and which it readily induces. The qualitative profile of responses to water blank stimuli has some similarity to that of D5′G and this confounds the determination of its psychophysical parameters. Its 50 per cent detection threshold is in the region of 0·005 per cent concentration.

It is recommended that all descriptions of the taste qualities of complex substances be expressed not just as mixtures of the four taste primaries, but as differences from the profiles generated by gustatory noise, where this is defined as a response to a blank stimulus in the experimental context considered. In this manner a taste may be characterized in two ways, as a change in the non-null profile proportions, or as a change in the probability of giving a non-null response but with the same profile proportions. The problem awaits re-examination in terms of subjects' individual differences in response profiles.

References

Adachi, A., Kawamura, Y., Ohara, M. & Ikeda, S. (1965). Neurophysical studies on taste effectiveness of sodium 5′ guanylate as a chemical taste enhancer. *Amino Acid Nucleic Acid, Tokyo* **12**, 63–68.

Ambos, M., Leavitt, N. E., Marmorek, L. & Wolschina, S. B. Sin Cib-Syn (1968). Accent on glutamate. *New Engl. J. Med.* **279**, 105.

Amerine, M. A., Pangborn, R. M. & Roessler, E. B. (1965). *Principles of Sensory Evaluation of Food.* New York and London: Academic Press.

Crocker, E. C. & Henderson, L. (1932). The glutamic taste. *Am. Perfumer Ess. Oil Rev.* **27**, 156–158.

Crocker, E. C. & Sjöström, L. B. (1948). Observational study of monosodium glutamate. *Fd Res.* **13**, 450–455.

Eberhardt, W. (1964). 5′-Ribonucleotide, new taste ingredients for the food industry. *Dt. LebensmittRdsch.* **60**, 102–103.

Good, I. J. (1965). *The Estimation of Probabilities: an Essay on Modern Bayesian Methods.* Cambridge, Mass.: M.I.T. Press.

Gregson, R. A. M. (1964). Modification of perceived relative intensities of acid tastes by ambient illumination changes. *Aust. J. Psychol.* **16**, 190–199.

Gregson, R. A. M. (1965). The effect of disodium inosinate on the perception of relative sweetness. (Research Project No. 8. Department of Psychology and Sociology, University of Canterbury, New Zealand.)

Gregson, R. A. M. (1969). Determination of a gustatory L70 point for sodium benzoate. *Br. J. Psychol.* **60**, 187–197.

Konusu, S., Maeda, Y. & Fijita, T. (1960). Inosinic acid and free amino acids in the flavouring substances in the Katsuobishi stock. *Nippon Suisan Gakkaishi* **26**, 45–48.

KUMINAKA, A., KIBI, M. & SAKAGUCHI, K. (1964). History and development of flavour nucleotides. *Fd Technol.* **18**, 30–35.

KWOK, R. H. M. (1968). The Chinese restaurant syndrome. *New Engl. J. Med.* **278**, 796

LOCKHART, E. E. & GAINER, J. M. (1950). Effect of monosodium glutamate on taste of pure sucrose and sodium chloride. *Fd Res.* **15**, 459–464.

MOSEL, J. N. & KANTROWITZ, G. (1952). The effect of monosodium glutamate on acuity to the primary tastes. *Am. J. Psychol.* **65**, 573–579.

PENAU, H. (1966). The use of inosinic and guanylic acids (sodium and calcium salts) in food. *Bull. Acad. natn. Méd.* **150**, 23, 435–442.

PILGRIM, F. J., SCHUTZ, H. G. & PERYAM, D. R. (1955). Influence of monosodium glutamate on taste perception. *Fd Res.* **20**, 310–314.

SATO, M., YAMASHITA, S. & OGAWA, H. (1966). Electrophysiological studies on the effects of ribonucleotides on taste receptors of rats: analysis of single unit response. *Amino Acid Nucleic Acid, Tokyo* **13**, 62–69.

SHIMAZONO, H. (1964). Distribution of 5′ ribonucleotides in foods and their application to foods. *Fd Technol.* **18**, 36–40.

TSENG, C. & CHU, E. J. (1933). Tastes of glutamic acid and related compounds. II. Tastes of some derivatives of *dl* glutamic acid. *J. Chin. chem. Soc.* **1**, 188–198.

VAN COFF, H., HAMILTON, C. E. & LITTELL, A. (1954). The effects of subthreshold concentrations of monosodium glutamate on absolute taste thresholds (Paper presented at 75th annual meeting of Eastern Psychological Association of New York.)

VAN SKRAMLIK, E. (1926). *Die Physiologie des Geruchs- und Geschmackssinnes.* Leipzig: Thieme.

YAMAGUCHI, S. (1965). Synergistic action of flavouring materials. *Shinku Kagaku* **13**, 101–105. (Japanese.)

YOSHIDA, M. (1963). Similarity among different kinds of taste near the threshold concentration. *Jap. J. Psychol.* **34**, 25–35.

YOSHIDA, M., NINOMIYA, T., IKEDA, S., YAMAGUCHI, D., YOSHIKAWA, T. & OHARA, M. (1966). Studies on the taste of amino acids. Determination of threshold values. *J. agric. chem. Soc. Japan* **40**, 295–299. (In Japanese with English summary.)

YOSHIDA, M. & SAITO, S. (1969). Multidimensional scaling of the taste of amino acids. *Jap. Psychol. Res.* **11**, 139–166.

(*Manuscript received* 10 *September* 1969; *revised manuscript received* 23 *January* 1970)

Br. J. Psychol. (1971), 62, 1, pp. 89-96
Printed in Great Britain

ORIENTATING RESPONSES OF MENTALLY RETARDED AND NORMAL SUBJECTS TO WORD-SIGNALS

By J. P. DAS and A. C. BOWER

Centre for the Study of Mental Retardation, University of Alberta, Edmonton, Canada

[illegible] gruent with an attentional deficit model of retardation.

Although several researchers (Zeaman & House, 1963; Ellis, 1963) have assumed that the retardate suffers from an attentional deficit, the exact import of this on a basic component of attention, the orientating response (OR), has not been clarified. The few studies in this area (Berkson *et al.*, 1961; Wolfensberger & O'Connor, 1965; Pilgrim *et al.*, 1969) have dealt with the habituation of OR to non-verbal stimuli such as tones or light flashes in order to examine whether or not retardates and normals differ in their rate of habituation. However, no clear answer emerges from these investigations.

The present study is concerned with the conditioning and habituation of ORs to meaningful verbal stimuli, some of which are specified as signals in a vigilance-type task. The index of OR is the occurrence of a galvanic skin response (GSR). GSRs may be evoked by non-critical stimuli during a general state of alertness, as well as by specific acts of attention to a relevant stimulus (Sokolov, 1960). Both may be observed in an experimental situation involving conditioning and habituation. We propose to examine how retardates may differ from normal children in these two processes. Words rather than light flashes are used here as signals to ensure the involvement of the verbal system. If retardates are deficient in verbal mediation or cannot maintain an instructed set (Luria, 1963), it should appear in the present task, as described in a later section. For it seems that, where non-verbal stimuli have been used, little difference between retardates and normals in habituation was noticed (cf. Wolfensberger & O'Connor, 1965).

Method

Subjects. Twenty-five retarded and 26 normal children served as subjects. The retardates came from a special school in the city and had IQs estimated to be from 40 to 65. Their CAs ranged from 13 to 16 years. The normals were grade 8 children, aged 13 and 14 years, from a nearby city school. Their IQ was 118·85 with an S.D. of 6·81. A number of subjects were excluded from the experiment. Twelve were retarded subjects unable to comprehend the instructions, as indicated by a failure to press the button to the imperative signal; also, nine retarded and 10 normal subjects who did not give 10 or more acceptable GSRs to the warning and/or the

imperative signal in the first 10 min., i.e. 50 per cent responding, were excluded from the analysis.

Task and procedure. The task involved listening for 30 min. to a series of six familiar words (box, man, cat, door, key, bell) presented by a tape-recorder in order to detect the occurrence of the signal word, 'man' (imperative signal). The signal 'man' always followed the word 'box' (warning signal). The sequence box–man occurred once during each minute for a total of 30 times throughout the experiment. An interval of 10 sec. separated each word. The position of the six words was randomly determined for each 1-min. period, with the exception that 'man' always followed 'box' and no word followed itself.

Responses were recorded with a multichannel Beckman polygraph. Two channels were used: one to measure the GSR, and the other to record the audio-signals from the tape-recorder and also to indicate the subject's button-pressing. The paper ran at a constant speed of 5 mm per sec., thus making it possible to determine the temporal location of each signal and button-press.

Zinc electrodes, 0·50 in. in diameter, were attached to the subject's thumb and wrist. Direct contact with the skin was avoided by mounting them on corn pads and injecting zinc sulphate paste between the skin and the electrode. The electrodes were secured with plastic electrical tape so as to minimize movement and pressure artifacts, but were not tight enough to obstruct blood circulation in that area of the skin (Brown, 1967, p. 18). Lykken's (1959) directions for making the electrodes and preparing the paste were followed.

Throughout the experiment, the subject sat in a large padded armchair with his arms resting on the chair's wide arms. The electrodes were placed on the left hand, leaving the right free to press the response button mounted on the right arm of the chair. The polygraph and tape-recorder were shielded from the subject by a partition. A one-way vision mirror in the partition allowed the experimenter to observe the subject, who sat with his back to the mirror. The subject was surrounded by bare walls in a space $5 \times 10 \times 8$ ft. high. The room temperature was thermostatically maintained at 72° F. The experimenter informed the subject that he was to listen to a number of words and press the button whenever he heard the word 'man'. The subject was told that 'man' would always follow 'box'. Once the subject had verbally demonstrated that he understood the instructions, the experimenter retired behind the partition and started the tape-recorder and polygraph.

Scoring. A GSR was identified as any downward pen deflexion of 1 mm or more in amplitude that occurred between 1 and 10 sec. after 'box' or 'man'. If the response occurred within 5 sec. of either word it was scored as a phase I GSR, and if it fell in the interval from 5 to 10 sec. after the word it was scored as a phase II GSR. This distinction was made following the work of Grings *et al.* (1962) and Prokasy & Ebel (1967). A GSR occurring in phase I after either should be an OR to that word. The phase II response would include delayed ORs to the word, which are expected to be very few, but in the case of 'box' the second response will also include anticipatory responses to 'man'.

Two frequency and latency measures were obtained. A count of all scorable GSR responses to 'box' and 'man' was one frequency measure, while all other GSRs (no temporal restrictions) were taken to be spontaneous responses. The first latency score (Lat_1) was the time from stimulus onset to the evocation of the GSR, while the second latency score (Lat_2) was represented by the time from the evocation to the peak of the GSR. Both measures have been frequently used by other investigators (cf. Clausen & Karrer, 1969).

It will be noted that no amplitude measure was calculated in the present study. This was done in our earlier study (Das *et al.*, 1969) where frequency, and not amplitude, reflected treatment differences. Other investigators have noticed this (Burnstein & Epstein, 1968) when the subject's record contained many missing GSRs to signals as in our experiments. Therefore, for the purposes of this study, only the frequency measures are presented.

Results

GSR frequencies to the signal and the non-signal words, and latencies of GSRs to signal words, were obtained for each subject. As mentioned earlier, the GSRs to signal words 'box' and 'man' were counted for two temporal phases. The four consecutive phases are labelled as Box I, Box II, Man I and Man II.

An analysis of variance on the GSR frequencies for signal words was performed, as shown in Table 1. ($n = 25$ in each group for all analyses of variance; thus one subject from the normal group was arbitrarily excluded.) It has one independent measure, the two subject *groups* (normal and retardates), and three repeated measures, which are: the two *words* ('box' and 'man') and two *positions* (phase I and II), and six *blocks* (5-min. periods into which the 30 min. task was divided for scoring). Table 1 shows no overall group differences in the total number of GSRs to the signal words

Table 1. *Analysis of variance for GSR frequencies*

Source	D.F.	M.S.	*F*	*P*
Groups (G)	1	2·96	< 1	n.s.
Error	48	9·87		
Words (W)	1	0·083	< 1	n.s.
G × W	1	19·76	12·19	< 0·01
Error	48	1·62		
Position (P)	1	1019·4	190·47	< 0·01
G × P	1	14·03	2·63	n.s.
Error	48	5·35		
Blocks (B)	5	8·76	6·80	< 0·01
G × B	5	2·96	2·30	< 0·05
Error	240	1·29		
W × P	1	35·36	13·49	< 0·01
G × W × P	1	2·43	< 1	n.s.
Error	48	2·62		
W × B	5	1·25	1·43	n.s.
G × W × B	5	0·37	< 1	n.s.
Error	240	0·88		
P × B	5	7·12	5·04	< 0·01
G × P × B	5	0·83	< 1	n.s.
Error	240	1·41		
W × P × B	5	0·48	< 1	n.s.
G × W × P × B	5	1·34	1·28	n.s.
Error	240	1·05		

for the 30-min. period. But there are group differences in responses to words, and for blocks. The significant groups × word interaction is caused by normals giving more responses to 'box' (mean 2·30) than to 'man' (mean 2·06), whereas the retardates give more responses to 'man' (mean 2·42) than to 'box' (mean 2·15). Thus the warning signal 'box' evoked ORs more frequently than 'man', which signalled button-pressing in the normal group, but the reverse was true for the retardates. The significant position effect which was obtained simply reflects the fact that mean GSRs to both 'box' and 'man' were more frequent during the first 5-sec. period than in the subsequent 5-sec. period.

Fig. 1 shows the GSRs to 'box' and 'man' for the two positions. The groups × word effect is clearly seen by considering Box I and Man I plots in this figure. It also shows that responses to Box II were more frequent than to Man II, giving rise to a significant word × position interaction. This is of special interest, for it confirms that anticipatory ORs to 'man' would enhance the GSR during the second phase of 'box', but, since no signal follows 'man', GSRs in the second phase of 'man' would be minimal.

Both groups showed a consistent decrement in GSRs over the last but one block

of 5-min. period, as seen in Fig. 2. In the last block the normals suddenly showed an increment in their GSR frequency, whereas the retardates continued to show a decrement. This resulted in a significant groups × block interaction. But the gradual decrement over the first five blocks was still strong enough to show a highly significant main effect for blocks.

The block × position effect was also highly significant. A close examination of the means revealed that the GSRs decreased gradually only for the first 5-sec. period (position I) following the signal words. This would be expected as ORs to the words are most likely to occur during this period.

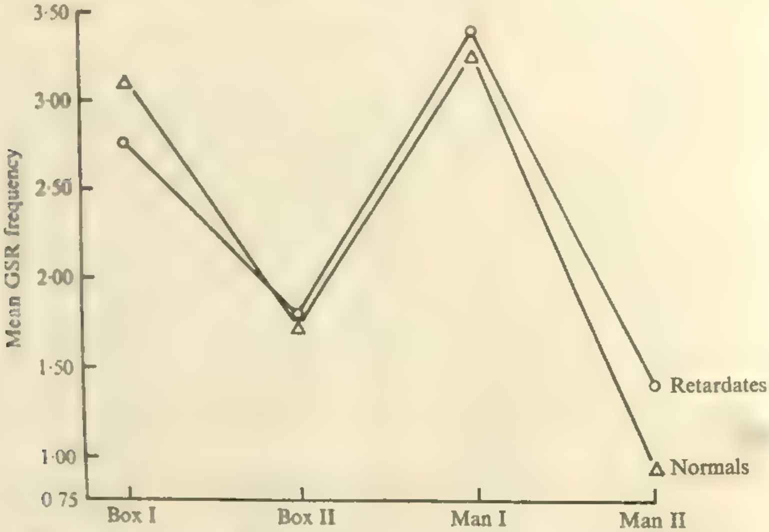


Fig. 1. GSRs to the warning ('box') and imperative ('man') signals during the first and second 5-sec. periods following signal presentation.

All other GSRs besides those to the signal words were similarly counted for blocks of 5 min., and were labelled as spontaneous GSRs. Fig. 3 shows the means for the two groups. The curves for the groups have almost the same shape, gradually descending up to the third block, followed by random fluctuations over the next three blocks. The retardates appear to have distinctly more frequent responses at each point compared to the normals. An analysis of variance for groups (2) and blocks (6) confirmed these trends to be significant: the main effect for groups had an F ratio of 12·75 (d.f. = 1, 48; $P < 0{\cdot}01$); the same for blocks had an F of 7·30 (d.f. = 5, 240; $P < 0{\cdot}01$). The interaction of groups × blocks approaches the 5 per cent level ($F = 2{\cdot}17$; d.f. = 5, 240; $P = 0{\cdot}06$), but does not appear to be clearly interpretable from an examination of the graphs in Fig. 3, except that the decrement from the first to the second block in the retardates is much more than the comparable decrement in normals.

Latency of a GSR has two components: the latency for evocation of the response (L_1) and the duration of the response (L_2). Latency scores were analysed by a 2 (groups) × 2 (words) × 2 (L_1, L_2) analysis of variance. F ratios for words (5·50; d.f. = 1, 48) and for group × word interaction (5·99; d.f. = 1, 48) were significant

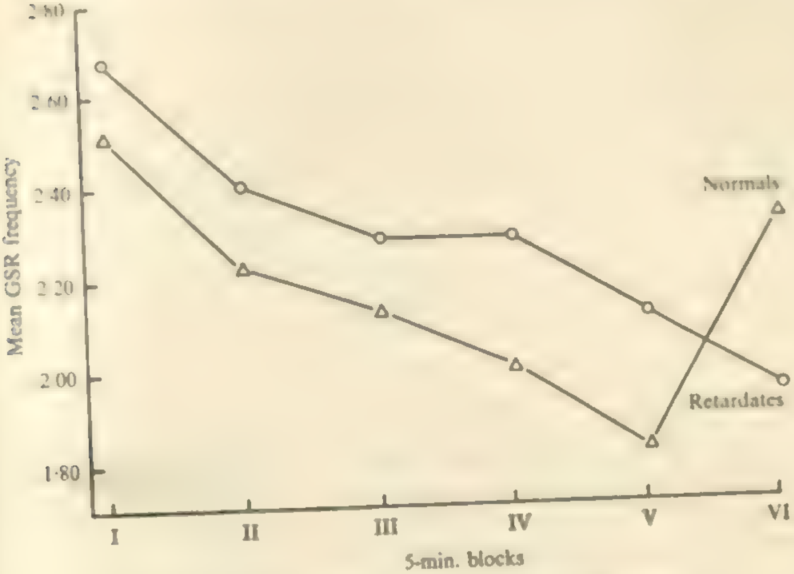


Fig. 2. Changes in GSR frequencies to signals over 30 min.

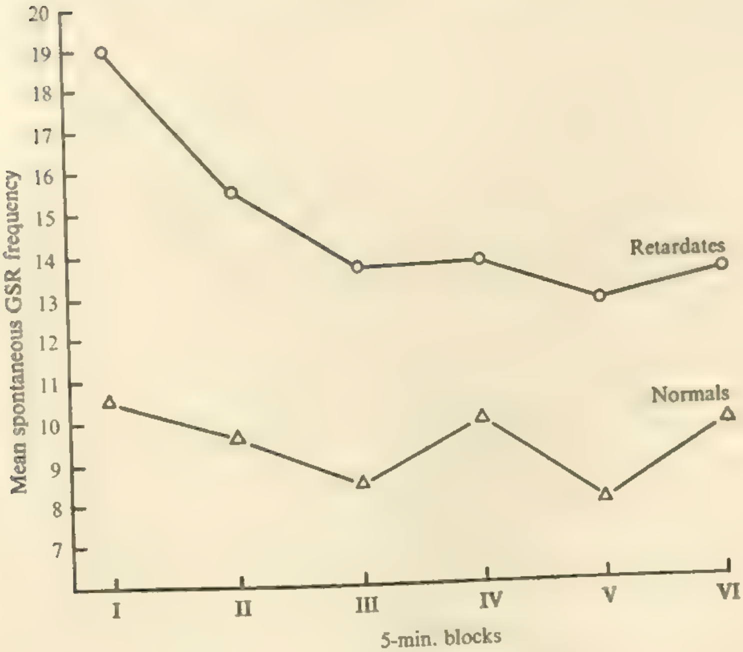


Fig. 3. Changes in spontaneous GSRs over 30 min.

below 0·025, and the F for group × latencies interactions ($F = 3{\cdot}25$; d.f. = 1, 48) approached significance ($P < 0{\cdot}10$). An examination of the means showed that latencies for evocation and duration were longer for the word 'box' than for 'man' in the normal group, but there was no difference between the words for the retarded group. Further, the group × latency interaction approached significance because the evocation latency (L_1) for 'box' was longer than for 'man' in the normal group, and, between the groups, it was longer in the normal than in the retarded. This is of some importance in that it reflects a holding back of the response to the warning signal 'box' on the part of the normals, an effect seen in delayed conditioned responses. The lack of an overall group difference in latencies argues against any slowing down of autonomic reaction time for the retarded subjects.

The motor response required of the subject was button-pressing upon hearing the word 'man'. On the whole the retardates were less accurate on this than the normals. They forgot to press the button more often than the normals, but they also made more false detections than the normals. None of the 26 normals had more than one omission during the 30-min. task, whereas seven of the 25 retardates had such errors. Similarly, none of the normals had more than one false detection as compared to 13 retardates. All of these differences yielded χ^2 values significant below 0·01. As we have noted before, the retardates gave more GSRs for 'man', the signal for button-pressing, than the normals, whereas they omitted the motor response more often than the normals. This suggests that the autonomic response to the signal was present without the accompanying motor responses in the retarded group. However, the opposite would have been expected from existing research and theory (Luria, 1963; Zeaman & House, 1963). Incidentally, the presence of GSR to signals without button-pressing and the fact that the normals had more button-pressing responses, but fewer GSRs following 'man', show that OR to 'man' could not be a consequence of the motor act of button-pressing.

Correlations between the frequencies of spontaneous and signal GSRs were computed in order to consider if individual differences in signal GSRs merely reflected those in spontaneous GSRs. The product-moment r between spontaneous GSRs and Box I GSRs was 0·26 ($P = 0{\cdot}07$). Similarly, for Box II, Man I and Man II, the r values were 0·40 ($P < 0{\cdot}01$), 0·27 ($P = 0{\cdot}06$) and 0·51 ($P < 0{\cdot}001$). All P values are based on a two-tailed test. It appears that GSRs in the second 5-sec. period following the signal were somewhat influenced by the condition which determined the frequency of spontaneous GSRs. But they are not merely an artifact of the level of spontaneous GSRs. The correlation between spontaneous frequencies of adjacent blocks should be lower than that between spontaneous and signal frequencies in the same block if the incidence of spontaneous GSRs completely determined that of the signal GSRs. However, the average correlation between adjacent blocks of spontaneous frequencies was 0·73, which was considerably higher than the highest correlation between signal and spontaneous GSRs we obtained (Man II, 0·51).

The relation between phase I and II frequencies was examined: r's between Box I and II, and Man I and II, were both 0·19 ($n = 50$, not significant). On the other hand, r between Box I and Man I was quite high (0·62). These results are interesting in the light of Prokasy & Ebel's (1967) observation that the phase I response is independent of II. Theoretically, phase I responses are ORs, and the high r suggests stable

individual differences in ORs for the preparatory and imperative signals. A statistically significant correlation between Box II and Man II ($r = 0{\cdot}43$) was also obtained, which can be partly accounted for by their common relationship with spontaneous GSRs.

Correlations between the frequency and latencies of signal GSRs were also computed, and reveal an inverse relationship: r between frequency and L_1 was $-0{\cdot}34$, frequency and L_2 was $-0{\cdot}37$. It implies a tendency in the subject who gives more frequent GSRs to react quickly to the signal, and have a GSR of shorter duration, compared to one who gives less frequent GSRs. There were no marked group differences in these correlations; if any, the retardates tended to give slightly higher correlations than the normals. The two latencies showed strong correlations with each other: r for the normals was 0·60, and for the retardates 0·61.

Discussion

OR specificity was clearly demonstrated in response to 'box'. As expected, the normals were more 'attentive' to the warning signal than the retardates. However, no gradual increment in ORs to 'box' was noticed in either of the two groups—both showed an overall adaptation. This was most likely a consequence of prior instruction to the subject that 'box' would invariably precede 'man'. As Fig. 1 shows, anticipatory OR to 'man' was in evidence for both groups of subjects during the second phase of 'box'. In habituation of specific OR, the two groups behaved indistinguishably except on the sixth block, where the normals showed a sudden rise in GSRs. We have no explanation for this, except that the normal subject might have picked up some cue as to the duration of the experiment. Since the entire vigilance task was taped, it is unlikely that such cues were present on the tape.

Rate of habituation of the non-specific OR, like that of specific OR, did not differentiate the two groups, as shown in Fig. 3, although the retardates were more sensitive to the non-signal stimuli. The same tendency is noticed even for signal stimuli, although the normal–retardate difference is not significant here. Their increased sensitivity or excitability is not maladaptive, however, as they habituate at the same rate as normals do. Thus our findings are not consonant with the general implications of an attentional deficit model of retardation in so far as these may relate to the ability of maintaining verbally mediated sets, or the retardates' marked distractability.

How similar was performance on this task to decrement in a vigilance task? Obviously, the present task of listening for 'man' was too simple to produce any gradual decrement in button-pressing. But, if we consider the autonomic rather than the motor component of vigilance behaviour, the adaptation of OR to the signals closely resembles a typical performance record in a 30-min. vigilance task (Das, 1964). We suggest that the OR measure of vigilance taps a basic response of the subject, and may be preferred to the conventional methods which require a motor response. Since the OR to the signal seems to precede the motor response, and can also be conceived as occurring earlier in the chain of neural events initiated by the signal, we suggest that the present method is specially appropriate for subject samples such as children and retardates with inadequate sensorimotor integration.

This research was partially supported by a National Research Council of Canada Grant APA 0292. The authors wish to thank the Winnifred Stewart School and the Edmonton Public School Board for their cooperation in providing the subjects.

REFERENCES

BERKSON, G., HERMELIN, B. & O'CONNOR, N. (1961). Physiological responses of normals and institutionalized mental defectives to repeated stimuli. *J. ment. Defic. Res.* **5**, 30–39.

BROWN, C. C. (1967). *Methods in Psychophysiology*. Baltimore: Williams & Wilkins.

BURSTEIN, K. R. & EPSTEIN, S. (1968). Procedure for reducing orienting reactions in GSR conditioning. *J. exp. Psychol.* **78**, 369–374.

CLAUSEN, J. & KARRER, R. (1969). Temporal factors in autonomic responses for normal and mentally defective subjects. *Am. J. ment. Defic.* **74**, 80–85.

DAS, J. P. (1964). Hypnosis, verbal satiation, vigilance and personality factors: a correlational study. *J. abnorm. soc. Psychol.* **68**, 72–78.

DAS, J. P., DYER, F. N. & BOWER, A. C. (1969). Orienting responses to vigilance signals: a comparison of GSRs of normal and retarded children. (Paper presented at XIX International Congress of Psychology, London.)

ELLIS, N. R. (ed.) (1963). *Handbook of Mental Deficiency*. New York: McGraw-Hill.

GRINGS, W. W., LOCKHART, R. A. & DAMERON, L. E. (1962). Conditioning autonomic responses of mentally subnormal individuals. *Psychol. Monogr.* **76**, no. 39.

LURIA, A. R. (1963). *The Mentally Retarded Child*. Oxford: Pergamon Press.

LYKKEN, D. T. (1959). Properties of electrodes used in electrodermal measurement. *J. comp. physiol. Psychol.* **52**, 629–634.

PILGRIM, D. L., MILLER, F. D. & COBB, H. V. (1969). GSR strength and habituation in normal and nonorganic mentally retarded children. *Am. J. ment. Defic.* **74**, 27–31.

PROKASY, W. F. & EBEL, H. C. (1967). Three components of the classically conditioned GSR in human subjects. *J. exp. Psychol.* **73**, 247–256.

SOKOLOV, E. N. (1960). Neuronal models and the orienting reflexes. In M. A. Brazier (ed.), *The Central Nervous System and Behavior*. New York: Josiah Macy Foundation.

WOLFENSBERGER, W. & O'CONNOR, N. (1965). Stimulus intensity and duration effects on EEG and GSR responses of normals and retardates. *Am. J. ment. Defic.* **70**, 21–37.

ZEAMAN, D. & HOUSE, B. J. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis (ed.), *Handbook of Mental Deficiency*. New York: McGraw-Hill.

(*Manuscript received* 23 *October* 1969; *revised manuscript received* 16 *December* 1969)

Br. J. Psychol. (1971), 62, 1, pp. 97–103
Printed in Great Britain

CONDITIONED EMOTIONAL RESPONSE PHENOMENA AND BRAIN STIMULATION

By JOHN T. HAWORTH
Department of Psychology, University of Manchester

Electrodes were implanted into the septal region in one group of rats and into the lateral hypothalamic region in another group. After stabilizing for lever pressing rates on a variable interval (VI) schedule for brain stimulation, each group was split into two. One group of septal and one group of hypothalamic animals were given the conditioning procedure for a CER whilst they could not obtain brain stimulation, and the other groups were given the conditioning procedure whilst they were lever pressing for brain stimulation. Manifestation of the CER was tested whilst the animals were bar pressing for brain stimulation and it was found possible to condition a CER in all of the animals. A second experiment, using one of the hypothalamic groups, showed that the CER was still manifest on a faster VI schedule, the same as that used for the septal animals. A third experiment, using one of the hypothalamic groups, showed that when the group was on a CRF schedule the CER was significantly attenuated even though it was later manifest on a VI schedule. This work indicates that the CER technique using brain stimulation reinforcement should prove useful for, amongst other things, the study of the interaction between motivation and conditioned emotional responses; also that comparisons between different brain loci for CER phenomena should be based on parametric studies of schedule of reinforcement and current intensity, etc.

The conditioned emotional response (CER) has often been used as a way of studying emotional and motivational behaviour in animals. The CER can be obtained in an animal to a previously neutral stimulus or conditioned stimulus (CS), as it is often called, such as a white light or clicker sound, by repeated presentations of CS–US pairings, where the US (the unconditioned stimulus) is, for example, a foot shock. This can be monitored visually and quantified by superimposing the CER on a baseline of operant responding by the animal when the CER is manifest as a drop in the rate of responding during the presentation of the CS. Using conventional reinforcement (liquid, solid food) for the baseline operant reinforcement, a CER has been obtained by numerous people (for example, Estes & Skinner, 1941; Hunt & Brady, 1951).

Few attempts have been made, however, to use this technique in conjunction with brain stimulation.

Brady & Conrad (1960) reported investigations into the CER with animals bar-pressing for high-value brain reinforcement (intracranial self-stimulation for electricity) on a VI_{60} schedule (variable interval with one reinforcement on average every 60 sec.). Three rats had single bipolar electrodes in the septal region of the brain and three had single bipolar electrodes in the medial forebrain bundle (MFB) anterior placements. After eight CER conditioning trials whilst the animals were bar-pressing for brain stimulation they could not obtain a CER, even though the animals would show a profound CER when given eight conditioning trials whilst bar-pressing for a conventional reward, in this case water. Brady (1961) also ran a study with rats on a VI_{15} schedule and obtained essentially the same results.

Brady & Conrad (1960) did, however, obtain a CER for brain stimulation after a large number of conditioning trials (23) at one of the septal placements in the albino

rat and also after only a few conditioning trials in two cats with an electrode in the head of the caudate nucleus. They suggested that their results could be explained on the basis of locus effects, i.e. whether or not brain stimulation attenuates a CER depends upon the locus at which the electrode is implanted in positive reinforcement areas of the brain. McIntyre (1966), using rats, has also obtained a CER whilst using high-value brain stimulation as baseline behaviour. His electrode placements though were restricted to the MFB lateral hypothalamic area.

Pliskoff *et al.* (1965) have pointed out, however, that under the appropriate experimental conditions brain reinforcement can be comparable to conventional reinforcement. Since a CER can readily be obtained using conventional reinforcement, then under some conditions of brain stimulation it could possibly be the case that a CER could be obtained with the electrode in any positive reinforcement area. Schwartzbaum & Donovick (1965) have also reported that in experiments where both peripheral shock and brain stimulation are used the outcome of the experiment could be a function of the degree of isolation between the brain-electrode circuit and the peripheral shock source. The possibility existed that the attenuation of the acquisition of a CER observed by Brady & Conrad could have been an artifact of experimental procedure. For these reasons the experiments of Brady & Conrad were repeated with design modifications and low-current values being used for the operant reinforcement.

METHOD

Subjects. The subjects were 16 male albino rats weighing 300–350 gm. Eight animals had one bipolar stainless steel electrode implanted into the lateral hypothalamic region. The coordinates, measuring from skull markings, were 4·5 mm posterior to bregma, 1·5 mm lateral to the mid-line and 8·5 mm below the surface of the skull. Eight animals had one bipolar stainless steel electrode implanted into the septal region, the coordinates being 1·5–3·0 mm anterior to bregma, 0·5 mm lateral from the mid-line and 5 mm below the surface of the skull.

Apparatus. The apparatus consisted of standard operant-conditioning equipment. The operant-conditioning box had a grid floor for shocking the animals and a small lever on one side to activate the brain-stimulation circuit. A neon house-light provided background illumination. A 7 W 110 V bulb placed just above the lever served as the CS. A similar 'shock box' without the lever served as the alternative conditioning chamber. The peripheral shock serving as the unconditioned stimulus was produced by a high resistance 60 c.p.s. source and was monitored across a 10 K resistor in series with the animal. The intracranial stimulation was provided by a constant current 60 c.p.s. power source, also monitored on a 10 K resistor in series with the animal.

Procedure. After a 1–3 week recovery period from the operation the animals were trained, using first of all a continuous and then a variable interval (VI) schedule of reinforcement, to stable rates of responding (lever-pressing) for brain stimulation. The stimulation was 60 c.p.s. sinusoidal ¼-sec. duration and of 30–65 μa. r.m.s. value.

For the hypothalamic animals stable rates of responding could be obtained on a VI_{60} schedule. For the septal animals, however, stable rates could only be maintained on a VI_{15} schedule. Two of the septal animals became diseased during this training period and had to be discarded.

The animals were pre-tested to the conditioned stimulus whilst they were lever-pressing until all of the animals showed no suppression of lever-pressing to the CS.

Conditioning was done in two ways. After splitting the hypothalamic animals into two groups of four and the septal animals into one group of four and one group of two animals, one hypothalamic and one septal group were conditioned for the CER in the operant-conditioning box whilst they were lever-pressing for brain stimulation on their appropriate VI schedules. These two groups were termed group II hypothalamic and group II septal respectively. The other hypothalamic and septal groups, termed group I hypothalamic and group I septal, were con-

Histology for a rat with an electrode in the septal region

JOHN T. HAWORTH

Histology for a rat with an electrode in the hypothalamic area

JOHN T. HAWORTH

ditioned in the alternative 'shock box' and the brain-stimulation leads were not clipped on to the electrodes. The animals thus had no access to intracranial self-stimulation. This method controlled for the possibility that intracranial self-stimulation may prevent the acquisition of a CER and also for the possibility that the conditioning shock to the feet might possibly induce a current in the brain-stimulation leads. Whether the animals were conditioned whilst obtaining brain stimulation or not, the detailed conditioning procedure was identical. Each animal had three conditioning trials during the 1 hr. experimental period. Each trial consisted of presenting the CS for 2 min. and terminating it with a shock delivered to the animal's feet. The intensity of the shock was 1·5 ma. r.m.s. value and the duration ½ sec. The procedure was repeated the next day.

On the third day each animal was tested for the CER response by presenting the CS alone three separate times during the 1 hr. period. The magnitude of the CER can be expressed by using a suppression ratio described by Annau & Kamin (1961). This is $B/(A+B)$, where B equals the number of bar-presses during the period of presentation of the CS (2 min.), and A = the number of bar-presses during the equivalent period immediately before CS presentation (2 min.). The ratio has limits of 0 and 1, where 0 indicates complete cessation of responding during the CS and 0·50 represents no effect of the CS. 1·00 is the case where there is no responding prior to the CS but where responding occurs during the CS. If the suppression ratio was not less than 0·5, the animals were given further conditioning trials.

Once the suppression ratio was less than 0·5, the animals were tested for 1 hr./day, with three CS presentations until the CER extinguished, or for a maximum of 5 days.

A second experiment was performed with the group II hypothalamic animals. The four animals of this group had not extinguished after 5 days' tests. After one week's rest they were switched from the VI_{60} schedule, which had been their schedule during Expt. I, to a VI_{15} schedule which had been the schedule for the septal animals during Expt. I. The hypothalamic animals were run for 3–4 days on VI_{15} until a stable rate of lever-pressing was obtained. They were then tested with the CS to see if the CER was manifest on the VI_{15} schedule.

A third experiment was done with the group II hypothalamic animals. The day following Expt. II they were tested with the CS whilst working on continuous brain reinforcement (CRF) as the initial schedule and then on VI_{15}.

RESULTS

At the end of the experiment the rats were sacrificed and the brains sectioned to determine the locus of the electrode tip. The histological results verify that the electrode tips reached the loci aimed for on implantation. Plate 1 shows an example of the septal histology and Plate 2 an example of hypothalamic histology.

The results for Expts. I and II are shown in Table 1. By testing for significance, using the Mann–Whitney U one-tailed test, between the mean suppression ratio for the last pre-test and the suppression ratio for the first presentation of the CS after the final shock, i.e. between columns 8 and 9 the CER in Expt. I is found to be significant at the 0·002 level. It was found possible to train all of the animals to manifest a profound CER. In Expt. II the CER is significant at the 0·01 level.

It is not really valid to make detailed comparisons between the hypothalamic and septal groups. Thus, even though statistics show that there is no significant difference between the CER for the hypothalamic group run at VI_{15} and the septal group run at VI_{15}, comparisons should not be forced between the groups since the hypothalamic group had a different experimental history from the septal group.

Comparisons made within the septal groups show that whilst it is possible to train groups to the same level of CER, the group being shocked whilst bar-pressing for brain stimulation needed a greater number of conditioning trials. Whilst it is not possible in this study to rule out any influence of the locus of implantation, it could well be that the brain stimulation or the peripheral shock interfered to some extent

with the acquisition of the CER, possibly by eliciting seizure activity. Bogacz *et al.* (1965) have shown that brain stimulation elicits seizure activity in septal animals. They have also shown that this is not the case for hypothalamic animals and the data in this experiment do not show a difference in the number of acquisition trials for the CER for the two hypothalamic groups. It must be borne in mind, however, that the hypothalamic group II animals were conditioned on a different schedule from that of the septal group II animals, and it may be that some other property of brain stimulation may have contributed to the resistance to acquisition in the septal animals.

Table 1. *CER data for Expts. I and II*

Group and locus	Rat no.	Schedule	Current intensity (μa.)	Mean total score* per hour	No. days conditioning	No. times shocked	Mean suppression ratio for last pre-test	Suppression ratio for first presentation of CS after final shock†
				EXPERIMENT I				
Hypothalamic group I	1488	VI_{60}	40	715	2	6	0·62	0·00
MFB lateral	1489	VI_{60}	30	613	2	6	0·45	0·00
hypothalamus	1491	VI_{60}	45	798	2	6	0·68	0·07
	1597	VI_{60}	55	1209	2	6	0·59	0·03
Hypothalamic	1596	VI_{60}	65	548	2	6	0·67	0·00
group II	1664	VI_{60}	60	1081	2	6	0·48	0·00
MFB lateral	1665	VI_{60}	60	471	2	6	0·47	0·67
hypothalamus	1666	VI_{60}	60	1964	2	6	0·57	0·16
Septal group I	1655	VI_{15}	45	402	2	6	0·53	0·00
Septal area	1657	VI_{15}	35	334	2	6	0·59	0·00
Septal group II	1693	VI_{15}	40	628	4	12	0·57	0·38
Septal area	1694	VI_{15}	35	375	3	9	0·57	0·00
	1281	VI_{15}	45	685	4	12	0·56	0·14
	1282	VI_{15}	35	953	2	6	0·57	0·18
				EXPERIMENT II				
Hypothalamic	1596	VI_{15}	45	708	2	6	0·67	0·00
group	1664	VI_{15}	30	1850	2	6	0·48	0·02
MFB lateral	1665	VI_{15}	35	1142	2	6	0·47	0·10
hypothalamus	1666	VI_{15}	30	2028	2	6	0·57	0·04

* For Expt. I the mean total score is the total of the scores/hr. for each experimental day from the day before pre-test to the final extinction of the CER, *or* day 5 of the extinction tests, divided by the number of days run. This gives an idea of the average rate of responding on the particular schedule of reinforcement concerned. In the case of Expt. II the mean total score is the total over 1 hr. for the last day of VI_{15} alone, i.e. before the CS is presented to test for CER on VI_{15}.

† For Expt. II this column refers to the first presentation of the CS after stabilizing runs on VI_{15} alone.

The results for Expt. III are given in Table 2 and in the cumulative record (Fig. 1).

Using the Mann–Whitney U one-tailed test there was no significant difference between the suppression ratio for the pre-test and the suppression ratio with the animals on continuous reinforcement. Brain stimulation on continuous reinforcement at these current values and at this locus apparently attenuates the CER. This result is not due to extinction of the CER, since a significant difference at the 0·01 level exists between the suppression ratio for the pre-test and the SR on VI_{15}. This difference supports the idea that the attenuation is a valid phenomenon. The schedule of

reinforcement is thus an important parameter at this locus and current values for the manifestation of a CER. These results have been replicated during part of a further study.

Table 2. *CER phenomena on different schedules of reinforcement: Expt. III*

(The animals were run on continuous reinforcement with one presentation of the CS, then switched without a break on to VI_{15} and the CS presented again. The same current intensity for the brain reinforcement was used for both schedules.)

Hypothalamic group II	Current intensity (μa.)	Score on CRF Pre-CS	Score on CRF CS	Score on VI_{15} Pre-CS	Score on VI_{15} CS	Mean SR for last pre-test	SR on CRF	SR on VI_{15}
Rat 1596	50	169	105	65	1	0·67	0·38	0·01
Rat 1664	35	144	152	41	1	0·48	0·51	0·02
Rat 1665	45	108	106	28	0	0·47	0·50	0·00
Rat 1666	30	158	165	48	20	0·57	0·51	0·29

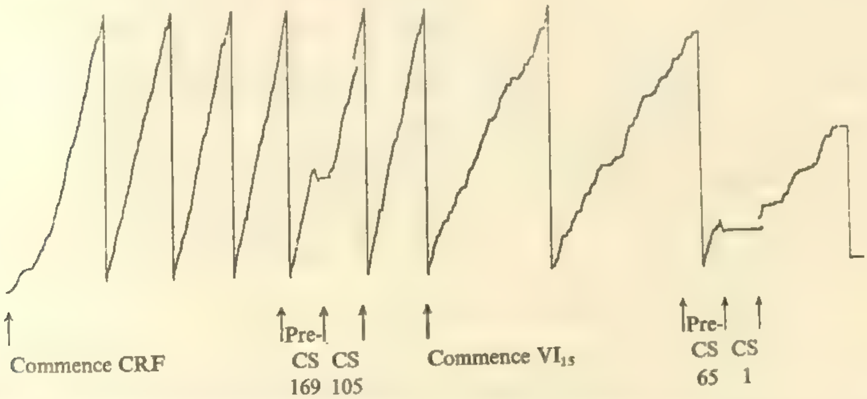


Fig. 1. Expt. III: cumulative record of CER phenomena on different schedules of reinforcement. Rat 1596. Locus: 4/1·5/8·5 MFB hypothalamus.

Discussion

The experiments have thus shown that it is possible to obtain a CER in rats using brain stimulation reinforcement at septal and MFB (hypothalamic) loci; also that a CER can be attenuated at hypothalamic loci and that the schedule of reinforcement is an important parameter in the manifestation of the CER at this locus. The septal rats in these experiments had electrode placements similar to the placements used by Brady & Conrad (1960). These results, taken together, indicate that the differences between the two experiments could be related to parameters of reinforcement rather than to simple locus effects. Distinct differences existed, for example, in the current intensity parameter of brain reinforcement. Brady & Conrad used a biphasic Lilly pulse of 0·2 msec. duration, 25 ma. peak, 100 c.p.s. and train duration of 0·5 sec. The peak of the sinusoidal pulse used in these experiments was only in the μa. range. It may well be that parametric differences in the characteristics of brain reinforcement can affect the manifestation and possibly the acquisition of a CER.

The results of Expt. III suggest that the attenuation of a CER may possibly be based on motivational factors. It may not be unreasonable to suppose that the manipulation of the schedules from VI to CRF resulted in increased motivation of the animal to attain brain reinforcement. The results of this experiment could thus be looked at in the form of an interaction between motivation and emotion, at higher motivational levels the conditioned emotional response being overcome. Study of the cumulative record for rat 1596 (Fig. 1) indicates that this rat initially suppressed slightly at the onset of the CS, but very quickly resumed self-stimulation during the remaining period of the CS. Similar traces have since been obtained. These results would not be inconsistent with a motivational interpretation. The evidence does contradict the lack of distractability hypothesis put forward by McIntyre (1966) to explain Brady & Conrad's findings of attenuation of the CER, since part of this hypothesis states that if an animal is once distracted by the CS it will not bar-press during the CS. McIntyre, in connexion with this hypothesis, also postulated that attenuation of suppression is limited to high-value brain stimulation.

The CER using brain stimulation for the positive reinforcement is a very flexible technique capable of a much finer manipulation than the CER using conventional reinforcement and it does not suffer from the same satiation effects. It may prove more useful than has done the CER using conventional reinforcement, for the screening of centrally acting drugs, for example.

It should also be fruitful to study the interaction between motivation and the CER by manipulating various parameters such as schedule of reinforcement and current intensity of brain stimulation affecting the motivational state of the animal and value of the US affecting the strength of the CER, suitable precautions being taken to control for unwanted locus effects.

Parametric studies of this nature could be done at different brain loci, allowing more valid comparisons to be made between different regions of the brain concerned in emotional/motivational behaviour. The US could also be delivered to negative reinforcement centres in the brain and information gained about the motivational interaction between various neural circuits such as the lateral hypothalamic system and the periventricular fibres described by Olds.

I am indebted to Dr Zoltan Annau, now at the Department of Environmental Medicine, Johns Hopkins University, U.S.A., for training and for the launching of this project.

REFERENCES

ANNAU, Z. & KAMIN, L. J. (1961). The conditioned emotional response as a function of the intensity of the US. *J. comp. physiol. Psychol.* **54**, 428–432.

BOGACZ, J., ST LAURENT, J. & OLDS, J. (1965). Dissociation of self-stimulation and epileptiform activity. *Electroenceph. clin. Neurophysiol.* **19**, 75–87.

BRADY, J. V. (1961). Motivational–emotional factors and self-stimulation. In D. E. Sheer (ed.), *Electrical Stimulation of the Brain.* Austin: University of Texas Press.

BRADY, J. V. & CONRAD, D. G. (1960). Some effects of limbic system self-stimulation upon conditioned emotional behaviour. *J. comp. physiol. Psychol.* **53**, 128–137.

ESTES, W. K. & SKINNER, B. F. (1941). Some quantitative properties of anxiety. *J. exp. Psychol.* **29**, 390–400.

HUNT, H. F. & BRADY, J. V. (1951). Some effects of electroconvulsive shock on a conditioned emotional response ('anxiety'). *J. comp. physiol. Psychol.* **44**, 88–98.

McIntyre, R. W. (1966). Conditioned suppression and self-stimulation. *Psychon. Sci.* **5**, 273–274.
Pliskoff, S. S., Wright, J. E. & Hawkins, I. D. (1965). Brain stimulation as a reinforcer: intermittent schedules. *J. exp. Analysis Behav.* **8**, 75–88.
Schwartzbaum, S. S. & Donovick, P. T. (1965). An artifact in the use of brain stimulation with shock motivated behaviour. *Psychon. Sci.* **2**, 183–184.

(*Manuscript received* 18 *December* 1969)

Br. J. Psychol. (1971), 62, 1, pp. 105–110
Printed in Great Britain

PREFERENCE FOR NON-OBJECTIVE ART: PERSONAL AND PSYCHOSOCIAL DETERMINERS*

By P. ROUBERTOUX, M. CARLIER AND J. CHAGUIBOFF
Institut d'esthétique et des sciences de l'art, University of Paris

The criteria of representational *v.* non-representational art were used to construct a scale of aesthetic preference which was found to be consistent and reliable. By the use of this scale, two groups of subjects were obtained: S_N who showed preference for non-representational art, and S_R who showed preference for representational art. S_N showed higher mean scores than S_R on the three scales of the 16 PF (dominance, dependent character, conservatism) and on the scale of globalism of approach. The socio-cultural background of the S_N's family was found to be significantly lower than that of S_R.

The purpose of this paper is to study the influence of personality variables and of differential variables on the preferences for aesthetic stimuli which differ in style (representational and non-representational paintings).

As stated by Knapp (1964), subjects preferring non-representational paintings 'might be described as subjective, imaginative, unpractical and sensitive' (p. 52). Psychosocial variables have also an influence, particularly the socio-economic status of the parents. But two critical remarks could be made.

(*a*) The authors ask the subjects to express their preferences by rating each stimulus picture. As Francès & Voillaume (1964) have shown, several criteria intervene in this kind of rating, and the criterion of spontaneous affective reaction generally used in studies on aesthetic preferences often suffers from lack of reliability in the ratings (Oléron, 1966). Underlying this criterion, there are several others which may be related to the form, colour, composition and themes of the paintings. A factor analysis by Guilford & Holley (1949) reveals the multiple nature of the criteria, and these authors report five factors: (1) special theme; (2) special design; (3) feminine taste; (4) luxury; (5) love of outdoor life.

Frequently personal and differential variables may be inserted within the S–R scheme. For example, Eysenck (1968) shows that introverted persons prefer narrow rectangles to wide ones. The Rorschach ranking test indicates that anxious and aggressive persons tend to reject cards saturated with black or red.

In the experiment by Knapp mentioned above, more or less pronounced realism plays a part among a set of uncontrolled variables. Formal stimuli (composition, drawing, colour) differ from one painting to another, in spite of the ratings of judges using approximative criteria. It is possible that those stimuli implicitly act so as to direct the preferences and rejections as well as the tolerances to less realism. There are two ways in which these interferences might be eliminated, i.e. either by restricting the criterion, or by more stringent control of the stimuli. The second alternative has been chosen here.

(*b*) In a previous study (Knapp & Wulff, 1963) but not in the later one (Knapp, 1964), the stimuli (which consisted of still-life paintings) were rated on a continuum

* Equipe no. 191 associée au C.N.R.S., Faculté des Lettres et Sciences Humaines de Paris-Ouest, Directeur de Recherche, Professeur R. Francès.

from the most realistic (17th and 18th centuries) to the least realistic (Cubist), through an intermediate (Impressionist) group. But Cubist painting is neither abstract nor representational. In the paintings chosen, a still life is pictured with less realism than in Impressionist painting and with even less realism than in 17th-century painting. What is measured here is not a preference for non-representational painting, but the amount of preferred verisimilitude between the object and its artistic representation. If it is assumed that figurative art is predominant in the subject's cultural background, what is measured may be taken to be simply the tolerance for representational distortion. The preference for an abstract painting compared to a figurative one thus implies tolerance of a lack of representation. This problem goes far beyond the field of experimental aesthetics and could be related to the general problem of tolerance of non-directivity. The aim of the present experiment is to determine the nature of those personality or differential variables which act to direct subjects' preferences for figurative or non-representational forms of art.

METHOD

Stimuli. The dependent variable consisted of the scores given on an art preference scale. This was derived from a set of 300 diapositives with an equal number of abstract and figurative paintings which could not be regarded as well known, but were painted by contemporary artists. Nudes and acts of violence were omitted. Twenty-four scales were constructed with the assistance of three judges. (One of the judges was the curator of a provincial museum with an important collection of contemporary paintings, and the other two were his assistants.) The purpose of the scales was to determine the formal qualities of the paintings. The scales were three-point scales, the end-points being defined by adjectives, e.g. colour: monochrome–contrasted; touch: heavy–light, etc. The criteria having been defined, the judges subsequently rated each painting on each of the 24 scales independently of each other. The paintings for which the judges' ratings were inconsistent were eliminated. By a simple sorting procedure it was then possible to match a painting of the abstract group with a painting of the figurative group which was identical, or approximately identical, with the first as to the formal qualities defined through the scales. Twenty-eight paintings* (i.e. 14 abstract–figurative pairs) were retained. The paintings in each pair were identical on at least 19 scales out of 24.

Personality scales and socio-cultural scales. The independent variables consisted of personality scales and psychosocial scales. The 16 Personality Factors of Cattell's questionnaire and the group version of the Zulliger test were used. Only the G, D and Dd types of approach were studied in this test. The procedure enabled us to construct, after standardization, a scale of the mode of approach of non-structured stimuli from the more global (high score) to the less global (low score).

Three scales of socio-cultural level were used: (1) the social status of the parents, which was calculated with the aid of socio-professional indices; (2) the educational level of the parents as measured by their university grades (the scores were weighted in order to yield a continuous scale 3 to 50; Zazzo & Hurtig, 1969); the cultural receptiveness of the subject as indicated by active or passive participation in any form of art (continuous scale from 1 to 20).

Subjects. Two hundred and forty-nine undergraduate students (20–23 years old) from the departments of the University of Paris–Ouest. All were volunteers and had been chosen at random on the campus. Art students were eliminated. There were 70 males and 179 females (this frequency is identical with that of the student population registered in this university).

The experiment. The paintings were shown simultaneously on two different screens, A and B, to groups of no more than five subjects at a time, who were instructed to respond by forced-choice technique. The following instruction was given: 'We are going to show you reproductions of paintings which were painted by contemporary artists; all of them are of good artistic quality and belong to national museums. You will see two works by each artist, painted at different

* The lists of scales and paintings are given by Roubertoux & Carlier (1969).

paintings, one being abstract and the other one representational. We shall show you the reproduction of the abstract painting simultaneously with the figurative one, on two screens. You will say which one you prefer by writing, on the sheet you have been given, the letter which designates the screen on which the preferred painting is shown. For instance, here are two works by the same painter who will be unnamed. If you prefer this one (A), write (A), if you prefer that one (B), write (B) on your sheet. There are neither correct nor incorrect answers. Please don't make any comment aloud.'

Some of the pairs which were presented were not painted by the same artist, but the instruction was intended to avoid a weighting on the subject's preference by the artist's prestige (Francès, 1963; Roubertoux, 1968). Furthermore, after the test the paintings were presented one by one and the subjects were told: 'If you know the name of the artist who painted such-and-such a painting, write it down in front of the corresponding number on the sheet.' The subjects who gave a name were eliminated in order to avoid the pseudo-prestige effect. In fact, we did not find such an effect, given the coherence of the items mentioned above; and it is to be noted that it could have influenced the preferences for the abstract paintings as well as for the figurative ones.

The preference scale for abstract art. The consistency of the 14 items of the scales was tested on 249 subjects. The percentages of preferences for abstract art varied between 41 and 62 per cent according to items. Kuder–Richardson's formula (21) gives a coefficient of internal consistency of 0·92 for the data. The scale may thus be considered homogeneous. A global score of preference for abstract art could therefore be computed, by addition of the item scores (if a non-representational painting was preferred, the answer was scored 1). The test–retest reliability was estimated for the choices, according to the following experimental design:

	Test	3 weeks	Retest
Test of preference for abstract art	(A)	(B . . . C)	(B+A+C)

B and C are series of items analogous to scale A. Only the correlation between results to scale A in both situations is taken into account. It was found to be 0·83 for randomly chosen subjects.

The preference scores were found to be normally distributed for 249 subjects. The confidence limits of the mean distribution were $6{\cdot}10 < m < 6{\cdot}82$. There was a very slight preference for figurative paintings. In order to eliminate effects of possible random choices due to the forced-choice technique, the method of contrasted groups was used. Deciles are drawn on the cumulative distribution curve. Only the two lower and the two upper classes are taken into account; they cover the following approximated scores of preference for abstract paintings. Group N, high score of preference for non-representational art ($n = 27$; 7 males, 20 females): scores 11, 12, 13, 14. Group R, low score of preference for non-representational art ($n = 27$; 10 males, 17 females): scores 1, 2, 3, 4. Each of the two groups was given the personality inventories and the social identification questionnaire.

RESULTS

No sex differences were found over groups N and R ($\chi^2 = 0{\cdot}78$). The mean scores of the two groups were compared by means of the t test for independent samples ($N_1 = N_2$) and the results are given in Table 1.

Referring to the graphs of Sakoda *et al.* (1954), chance probability of obtaining at least n' significant statistics at the 0·05 level from n calculated statistics is $0{\cdot}05 < P < 0{\cdot}01$.

DISCUSSION

Personality variables. The subjects who tolerate pictorial non-representation (S_N) and those who do not (S_R) are discriminated by three primary personality traits. In decreasing order of discriminatory power, they are: E, *dominance v. submissiveness*; G, *positive superego character v. dependent character*; Q_1, *radicalism v. conservatism*. The N and R groups also differ significantly regarding cognitive style, i.e. globalism of approach.

(a) Generally speaking, trait E indicates that S_R are less dominant than S_N. According to Cattell's questionnaire (1965), the S_N may be said to show greater ascendance, and to show more self-confidence. The S_R, on the other hand, are more submissive and more easily perturbed by unexpected situations. Subjects who tolerate pictorial non-representation are leaders; those who do not tolerate it are followers. It is as if S_R were more sensitive than S_N to the ambiguity of a non-structured pictorial stimulus, as if they needed a guiding scheme in the exploration of the painting; and the representation, as long as it introduces a guiding scheme, simplifies the stimulus by organizing it. For a non-trained subject, a non-representational pictorial stimulus is unorganized, and by this fact it has a high degree of

Table 1. *Average scores of the two groups (m_N, m_R) on the personality and social scales*

Scales	m_N	m_R	t
A affectothymia	4·88	4·51	0·54
B intelligence	6·11	5·55	0·98
C ego strength	3·92	3·18	1·13
E dominance	4·18	3·14	3·49***
F surgency	4·88	4·51	0·54
G superego strength	2·51	3·62	2·78***
H parmia	4·03	3·70	0·54
I tender-mindedness	6·40	5·88	0·91
L suspiciousness	5·29	5·33	0·00
M non-conformity	6·74	6·51	0·40
N shrewdness	4·18	4·66	0·83
O guilt-proneness	6·33	6·37	0·00
Q_1 liberalism	6·66	5·77	1·71*
Q_2 self-sufficiency	4·81	4·44	0·69
Q_3 self-sentiment	2·88	2·85	0·00
Q_4 ergic tension	5·85	6·11	1·25
Mode of approach of non-structured stimuli	18·00	20·51	2·06**
Parents' social status	25·33	32·29	4·08****
Parents' educational level	27·74	33·51	1·52
Cultural receptiveness of the subject	7·48	4·03	1·70

**** $P \leqslant 0·001$. *** $P \leqslant 0·01$. ** $P \leqslant 0·05$. * $P \leqslant 0·10$.

complexity and of ambiguity. Berlyne (1960) showed that collative variables (complexity, ambiguity) may direct aesthetic choice. But he also admits that personality traits are related to tolerance of complexity and ambiguity. The studies of Gardner and his collaborators are in the same trend (Gardner & Lohrenz, 1960; Gardner & Long, 1962). It is as if the subjects were in a condition of non-directiveness (in the sense of groups dynamics) when faced with an exploratory task.

(b) Trait G reveals a higher degree of dependency in S_N, S_R being more determined. This is not contradictory with the analysis of trait E, since dominance *v.* submissiveness is slightly correlated with trait G, and item-by-item analysis reveals a different content. S_N are changeable, fickle and have a liking for discontinuity and for innovation. Abstract art, for subjects who know little about art, has the connotation of novelty. Owing to their superficial attitude, they accept the lack of criteria. They can choose between the multiple attractions which the non-structured task elicits. Contrary to S_R, who are conscientious and use stable criteria, S_N are able to appreciate the multiple attractions of abstract painting. S_N are in opposition, and hence approve

the norms which are not approved of by their own in-group. Preference for abstract art, which is still considered avant-garde in some social classes, may be related to this trait. Finally, S_N are very fragmentary in their analyses.

(*c*) Factor Q_1, at a very low level of significance, links preference for abstract art and conservatism in a paradoxical way; and the questions which constitute the scale items are concerned with the subject's fundamental attitudes: political, religious, economic or ethical. Consequently, it is not possible to predict the attitudes towards new art from the attitudes towards new social ideas.

(*d*) Globalism in approach is empirically defined by the authors as an attitude of giving precedence to a general structure over a detailed structure in an initially non-structured set. These two types of approach to a complex stimulus discriminate groups N and R. S_N stick to the details, S_R to the whole. That was the prediction to which the interpretation of trait G gave rise: a strict organization of thought which tends to apply a certain model in a rigid manner. S_R are easily disturbed by an unexpected situation (E), especially when their organization of thought fails. Slightly more strained and irritable than S_N, S_R lack tolerance of frustration and thus are led to reject an ambiguous stimulus which they consider as frustrating.

(*e*) It was hypothesized that introversion might orientate the choice towards one form of art rather than another (Francès, 1968). This hypothesis was not verified either by the present experiment or by an earlier experiment by Roubertoux & Carlier (1969) in which several scales of introversion, dealing with several meanings of this term, did not present any correlation with scale of preference for abstract art which was used in the present study.

Psychosocial variables. The present result is completely inconsistent with those of different works in this field, a disagreement which might be explained by the fact that the groups differ (Knapp & Wulff, 1963). Though S_N show a lower educational level than S_R, the two groups do not differ significantly in this. But S_N who prefer abstract art come from a significantly lower socio-economic background than those who prefer figurative art. This fact may be explained by the hypothesis that subjects of group N (extracted from the less economically favoured group) consider interest in art, and especially preference for avant-garde art, as one of the norms of more socio-economic favoured groups.

But such conclusions can only be generalized very cautiously. The results have been obtained with subjects who have a low artistic knowledge (the subjects who identified the painters having been eliminated in order to avoid a prestige effect). A comparison of the relationships obtained with subjects having different degrees of artistic knowledge shows that the influence of personality variables decreases when artistic knowledge increases (Child, 1965). The main concern of our next study will be to see whether the observer phenomena are stable, whatever the subjects' degree of artistic knowledge may be.

References

Berlyne, D. E. (1960). *Conflict, Arousal and Curiosity.* New York: McGraw-Hill.

Cattell, R. B. (1965). *The Scientific Analysis of Personality,* Baltimore: Penguin Books.

Child, I. L. (1965). Personality correlates of aesthetic judgment in college students. *J. Personality* **33**, 476–511.

Eysenck, H. J. (1968). La personnalité et l'esthétique des formes simples. *Sciences de l'Art* **5**, 3–9.

FRANCÈS, R. (1963). Limite et nature des effets de prestige. II. Notoriété de l'auteur et jugement de l'oeuvre. *J. Psychol. norm. pathol.* **4**, 437–455.

FRANCÈS, R. (1968). *Psychologie de l'esthétique*. Paris: Presses Universitaires de France.

FRANCÈS, R. & VOILLAUME, H. (1964). Une composante du jugement pictural: la fidélité de la représentation. *Psychol. franç.* **9**, 241–256.

GARDNER, R. W. & LOHRENZ, L. J. (1960). Leveling-sharpening and serial reproduction of a story. *Bull. Menninger Clin.* **24**, 295–304.

GARDNER, R. W. & LONG, R. I. (1962). Control, defence and centration effect: a study of scanning behaviour. *Br. J. Psychol.* **53**, 129–140.

GUILFORD, J. P. & HOLLEY, J. W. (1949). A factorial approach to the analysis of variance in aesthetic judgments, *J. exp. Psychol.* **39**, 208–218.

KNAPP, R. H. (1964). An experimental study of triadic hypothesis concerning the sources of aesthetic imagery. *J. project. Tech.* **28**, 49–54.

KNAPP, R. H. & WULFF, A. (1963). Preferences for abstract and representational art. *J. soc. Psychol.* **60**, 255–262.

OLÉRON, G. (1966). Jugement esthétique et temps de décision. *Sciences de l'Art* **3**, 47–68.

ROUBERTOUX, P. (1968). Contribution des variables psycho-sociales à l'orientation des choix esthétiques spontanés (notoriété et communication persuasive). *Sciences de l'Art* **5**, 35–49.

ROUBERTOUX, P. & CARLIER, M. (1969). Une dimension de la personnalité—l'introversion et les préférences esthétiques. *Psychol. franç.* (in press).

SAKODA, J. M., COHEN, B. J. & BEALL, G. (1954). Test of significance for a series of statistical tests. *Psychol. Bull.* **51**, 172–175.

ZAZZO, B. & HURTIG, M. C. (1969). Milieu et côte sociale. In R. Zazzo (ed.), *Croissance de l'enfant, genèse de l'homme*. Paris: Presses Universitaires de France.

(*Manuscript received* 3 *December* 1969; *revised manuscript received* 9 *March* 1970)

Br. J. Psychol. (1971), 62, 1, pp. 111–112
Printed in Great Britain

HANDEDNESS AND THE PATTERN OF HUMAN ABILITY

By EDGAR MILLER
Department of Psychology, University of Hull

Levy (1969) has suggested that 'left-handed' persons have their verbal functions represented in both hemispheres of the brain whilst right handers have their verbal functions confined to the left hemisphere. She argued that the verbal functions undertaken by the right hemisphere in the left hander will be at the expense of the spatial functions normally subserved by that hemisphere. On testing 15 right handed and 10 left handed postgraduate students on the Wechsler Adult Intelligence Scale she found her expectations confirmed in that the right handed group had a significantly higher Performance (visuo-spatial) IQ than the left handed group, but the two groups did not differ in Verbal IQ.

Annett (1964) proposed a model of handedness with a genetic basis and confirmed this empirically (Annett, 1967). This model categorizes handedness into three groups: right-, mixed and left-handedness. She suggested that pure left-handedness is relatively rare and that in most descriptions of handedness the term 'left-handed' usually refers to the predominantly more numerous mixed-handers as well as to pure left handers. It is the mixed-hander in her system who has speech mediated by both hemispheres, the pure left-hander being right hemisphere dominant for speech. Annett (1964) further suggested, on the basis of evidence from studies of brain-damaged children, that the biologically more primary visuo-spatial skills are developed in preference to verbal abilities. Assuming the legitimacy of extrapolating findings from brain-damaged children to the normal adult, Annett's general theory would predict the opposite finding to Levy, i.e. that mixed ('left') handers should be similar to right-handers in visuo-spatial functioning but be relatively impaired verbally.

The aim of the present investigation was to gather further data relevant to this issue by means of a partial replication of Levy's experiment. It is assumed that Levy's 'left-handed' subjects were at least predominantly mixed-handed in Annett's terminology.

The subjects used were first-year psychology undergraduates who, at the time of testing, were ignorant of the experimental hypotheses. Handedness was assessed by the questionnaire given in Annett (1967), which revealed the sample to contain 29 right-handers and 23 mixed-handers (the results of two cases of pure left-handedness were excluded). The measures used were the NIIP Group Test 33 (a test of verbal intelligence) and the NIIP Form Relations Test (a test requiring the visual manipulation of shapes in both two and three dimensions). Both measures were given as group tests following the instructions in the manuals.

The results are shown in Table 1. It can be seen that the two groups have almost identical mean levels on the verbal test but differ on the spatial test, where the right-handers have a significantly higher mean score than the mixed-handers. The results therefore confirm Levy's (1969) finding.

It might be suggested, in partial defence of the hypothesis derived from Annett (1964), that differences in verbal intelligence do exist between right- and mixed-handers but that these have not appeared, due to the highly selected nature of the subjects used. Because students are highly selected through the educational system for verbal intelligence, groups of students should contain only mixed-handers of high verbal intelligence and this could mask a real difference in the population at large. The plausibility of this explanation is much weakened in the present investigation since Annett (1967) has shown that the proportion of mixed-handers in undergraduate populations is the same as that occurring in unselected secondary school children and a group of less educationally advantaged adults (enlisted men).

Table 1. *Mean scores of right- and mixed-handers*

Test	Handedness		Significance
	Right	Mixed	(t test)
NIIP Group Test 33	153·3	153·6	n.s.
NIIP Form Relations Test	41·4	35·4	0·025

It must be concluded that real differences in ability between right- and mixed-handers do exist and probably reflect underlying differences in the asymmetrical organization of functions within the brain. It would be expected that pure left-handers, who are presumably right speech dominant, would show a similar pattern of abilities to right-handers but this is difficult to test because such persons comprise only 4 per cent of the population.

REFERENCES

ANNETT, M. (1964). A model of the inheritance of handedness and cerebral dominance. *Nature, Lond.* **204**, 59–60.

ANNETT, M. (1967). The binomial distribution of right, mixed and left handedness. *Q. J. exp. Psychol.* **19**, 327–333.

LEVY, J. (1969). Possible basis for the evolution of lateral specialization of the human brain. *Nature, Lond.* **224**, 614–615.

(*Manuscript received* 7 *February* 1970)

Br. J. Psychol. (1971), 62, 1, *pp.* 113–117
Printed in Great Britain

A CLARIFICATION OF CROPLEY AND MASLANY'S ANALYSIS OF THE WALLACH–KOGAN CREATIVITY TESTS

By NATHAN KOGAN*

Educational Testing Service, Princeton, New Jersey, U.S.A.

In a recent article Cropley & Maslany (1969) examined the relationships among and between the Wallach–Kogan creativity tasks and the Thurstone Primary Mental Abilities tests in a sample of university students. A factor analysis yielded a general factor with moderate to substantial loadings for all of the tests in the battery. On this basis Cropley & Maslany concluded that complete statistical independence of the 'creativity' and 'intelligence' domains had not been established. The present paper attributes the Cropley–Maslany outcomes to a failure to rotate, and reports a Promax rotation of the Cropley–Maslany principal-components solution. Such rotation yields a pure 'creativity' factor, and two 'intelligence' factors. Unlike the Cropley–Maslany principal-axis factor loadings, the Promax solution is congruent with most of the evidence reported in other relevant investigations of the creativity–intelligence distinction.

Since the publication of the Wallach & Kogan (1965) monograph on the creativity–intelligence distinction, several reanalyses and partial replications have appeared in the research literature. Reanalysis has generally taken the form of a factor analysis of the creativity–intelligence correlations reported in the original Wallach–Kogan work. Two such analyses have been published (Fee, 1968; Ward, 1967), both supporting the Wallach–Kogan conclusion that 'creativity' and 'intelligence' become separate dimensions of cognitive functioning when divergent-thinking tasks are administered as games in a permissive testing context without time limits.

The sample in the Wallach–Kogan research consisted of children aged 10–11 enrolled in the fifth grade of an American elementary school. There have been several recent efforts to observe whether the Wallach–Kogan conclusions can be generalized to children of younger age, to high-school adolescents, and to university students. One must also consider the matter of replicability with children comparable in age to those employed in the Wallach–Kogan investigation. Successful replications of the foregoing sort have been reported by Pankove & Kogan (1968) and Kogan & Morgan (1969).

Where younger children are concerned, Ward (1968) adapted the Wallach–Kogan tasks for 7- and 8-year-olds and for kindergarten children aged 4–6. For the former sample, creativity and IQ were statistically independent; for the latter sample, independence was obtained only in the case of creativity tasks with semantic content. On the whole, Ward's results successfully extended the generalizability of the Wallach–Kogan research *downward* in age.

Age generalizability *upward* has also been demonstrated. Wallach & Wing (1969) observed that verbal and mathematical scores of the Scholastic Aptitude Test (College Entrance Examination Board, 1966) were unrelated to the fluency and uniqueness scores derived from the Wallach–Kogan creativity tasks. The sample in the foregoing research consisted of a select group of secondary-school students (in the age range 17–18) tested after notification of admission to a particular university. Studies based

* Present address: Graduate Faculty, New School for Social Research, New York, N.Y. 10011.

on university students in Australia (Cropley, 1968) and Canada (Cropley & Maslany, 1969) have yielded results largely consistent with the outcomes of the Wallach–Wing research. Regrettably, certain features of Cropley's work confuse the general pattern of findings described above, and hence provide the rationale for the present paper.

As the Cropley & Maslany (1969) study represents an elaboration and extension of earlier work by Cropley (1968), the 1969 work will serve as the object of critical attention. Given the similarity in substance, method and outcome between the two studies, however, many of the issues raised in the present paper will apply to both of Cropley's published articles.

Note should be taken of the difference in experimental and scoring procedures between the Cropley–Maslany study and the Wallach–Kogan research. Whereas the latter authors relied exclusively on individual administration of tasks, the former made use of a group administration format. Wallach & Kogan treated fluency and uniqueness as separate components of creativity, whereas Cropley & Maslany developed a single index of creativity combining the above components. The latter was accomplished by means of a weighting system in which higher or lower scores were assigned to responses on the basis of their statistical uncommonness. Finally, Wallach & Kogan employed the SCAT–STEP battery (Cooperative Test Division, 1957*a*, *b*) and selected subtests of the WISC (Wechsler, 1949) as 'intelligence' measures, whereas Cropley & Maslany used the test of Primary Mental Abilities (Thurstone, 1963). Despite these variations in tests and procedures across studies, the reliabilities and intercorrelations (reported in Table 1 of the Cropley–Maslany study) yielded a pattern corresponding quite closely to that found in the original Wallach–Kogan research.

This pattern is characterized by a strong trend towards significant positive correlations between tests within the 'creativity' and 'intelligence' domains, respectively, and negligible non-significant correlations across domains. On the basis of the correlational outcomes of the Cropley–Maslany study, it would appear that the Wallach–Kogan conclusions regarding the creativity–intelligence distinction are valid across a developmental span extending from the kindergarten to the university. The relationship, furthermore, seems to be a robust one in the sense that test and procedural modifications, which are often necessary to accommodate a diversity of age-groups, do not detract from the essential statistical independence of the 'creativity' and 'intelligence' measures. This is not intended to imply that any kind of variation in test content, scoring procedure or testing context is possible without altering the creativity–intelligence relationship. It should be noted that Cropley & Maslany, though employing group administrations, nevertheless maintained a testing atmosphere distinguished by informality, permissiveness, non-evaluation and the absence of time limits. Other research has shown that evaluative testing conditions alter the creativity–intelligence relationship in the direction of producing significant positive correlations (Boersma & O'Bryan, 1968). With respect to scoring procedures, Cropley & Maslany's decision to fuse the fluency and uniqueness dimensions into a single index would not be expected to affect the creativity–intelligence relationship given the high correlations usually found between the foregoing dimensions. Such correlations are consistent with an associative theory of creativity (e.g. Mednick, 1962). In sum, age generalizability of the Wallach–Kogan pattern of findings is obtained even

in the face of procedural variation. It must be noted, however, that where such variations violate the theoretical guiding principles governing the original Wallach–Kogan research, replicability and generalizability can no longer be expected (Kogan, 1969).

The congruity of the Cropley–Maslany correlational outcomes with other relevant research contrasts with the incongruity posed by those authors' factor analytic outcomes (as reported in Table 2 of their article). A strong general factor was obtained with moderate to substantial loadings on all of the tests included in the study. The second factor was bipolar, 'creativity' tests at one pole and 'intelligence' tests at the other. The third factor pointed to bipolarity in the 'intelligence' domain, verbal aptitude contrasting with numerical–spatial abilities. Cropley & Maslany concluded that the Wallach–Kogan 'tests measure a stable and internally consistent intellective mode, albeit one which is substantially related to general intelligence' (p. 398). In the present author's view, the foregoing statement confuses the issue. The fundamental purpose of the Wallach–Kogan research was to determine whether creativity and intelligence could be differentiated as independent dimensions when test content and context met certain clearly defined specifications. The outcomes of that research and most of the research that has followed (including the correlational portion of the Cropley–Maslany study) have supported complete independence between the two basic modes of thinking at issue. Since the Cropley–Maslany factor analysis does not show the kind of statistical separation expected, it behooves us to examine that analysis quite carefully.

Evidence for a general factor was obtained even though the highest correlation coefficient between a 'creativity' and an 'intelligence' test was 0·20, and the mean coefficient for all such cross-correlations was approximately 0·07. It should be further noted that factors II and III were bipolar (with negative loadings in the 40s), despite the fact that the highest negative correlation coefficient was −0·07. It is exceedingly unusual to obtain bipolar factors in the ability domain.

In the present author's judgement the difficulty of interpretation posed by the Cropley–Maslany factor-analytic outcomes derives from a failure to rotate their principal components solution. It is difficult to comprehend why no such rotation to simple structure was attempted. Both the Ward (1967) and Fee (1968) factor analyses of the original Wallach–Kogan data employed rotation. Conceivably, Cropley & Maslany might have been concerned that manual rotation opened the possibility of biasing the outcomes in the direction of the Wallach–Kogan hypotheses. With the ready availability of analytic rotational programmes, however, the problem of bias poses no danger.

Table 1 presents the factor loadings generated by a Promax (Hendrickson & White, 1964) rotation of Cropley & Maslany's principal components solution. The results are strikingly clear. Factor I is an unequivocal creativity factor. None of the Thurstone PMA tests yield meaningful loadings on factor I. The 'intelligence' tests show less homogeneity, four of the PMA tests loading heavily on factor II and the remaining two yielding high loadings on factor III. The numerical and spatial tests are substantially loaded on factor II, but 'letter series' is also part of the factor. The PMA tests with the highest loadings on factor III are of a verbal–semantic character. Factor loadings in the 30s for two of the creativity tasks on factor III do not seem high

enough (relative to the differential magnitude of loadings observed) to alter the general pattern of a creativity–intelligence separation.

The correlations between the Promax factors (Table 2) offer further support for the clear statistical separation of the 'creativity' and 'intelligence' domains. Factor I—the 'creativity' factor—is essentially uncorrelated with either 'intelligence' factor. Factors II and III, on the other hand, do correlate with one another to a slight degree.

Table 1. *Factor structure from Promax rotation of principal axes provided by Cropley & Maslany* (1969)

(All decimal points omitted. Variables 1–5 are the Wallach–Kogan creativity tasks; variables 6–11 are the Thurstone PMA tests.)

	Factors		
Variables	I	II	III
1. Names	627	−068	350
2. Uses	810	−049	272
3. Similarities	801	−147	111
4. Pattern meanings	803	123	248
5. Line meanings	752	128	341
6. Verbal meaning	−007	−035	815
7. Number facility	−036	602	219
8. Letter series	−082	549	288
9. Word grouping	056	235	663
10. Number series	−066	743	−053
11. Spatial relations	128	694	−081

Table 2. *Intercorrelations of Table* 1 *factors*

	I	II	III
I	1·000	0·124	0·068
II	0·124	1·000	0·286
III	0·068	0·286	1·000

In sum, rotation of Cropley & Maslany's principal axis factors clarifies the interpretative difficulties raised by their work and hence renders it consistent with the accumulated body of empirical evidence described in the present paper. At both the correlational and the factor-analytic level, the Wallach–Kogan tasks, when administered under permissive conditions, stand apart from the tests of convergent thinking that constitute the traditional 'intelligence' domain.

The present work was supported by the National Institute of Child Health and Human Development under Research Grant 1 PO1 HD-01762. I am grateful to Henrietta Gallagher for her assistance in statistical analysis and to William C. Ward for comments on the manuscript.

REFERENCES

BOERSMA, F. J. & O'BRYAN, K. (1968). An investigation of the relationship between creativity and intelligence under two conditions of testing. *J. Personality* **36**, 341–348.

COLLEGE ENTRANCE EXAMINATION BOARD (1966). *College Board Score Reports: A Guide for Counselors and Admissions Officers.* New York.

COOPERATIVE TEST DIVISION (1957*a*). *Cooperative School and College Ability Tests: Technical Report.* Princeton, N.J.: Educational Testing Service.

Cooperative Test Division (1957*b*). *Cooperative Sequential Tests of Educational Progress: Technical Report*. Princeton, N.J.: Educational Testing Service.

Cropley, A. J. (1968). A note on the Wallach–Kogan tests of creativity. *Br. J. educ. Psychol.* **38**, 197–201.

Cropley, A. J. & Maslany, G. W. (1969). Reliability and factorial validity of the Wallach–Kogan creativity tests. *Br. J. Psychol.* **60**, 395–398.

Fee, F. (1968). An alternative to Ward's factor analysis of Wallach and Kogan's 'creativity' correlations. *Br. J. educ. Psychol.* **38**, 319–321.

Hendrickson, A. E. & White, P. O. (1964). Promax: a quick method for rotation to oblique simple structure. *Br. J. statist. Psychol.* **17**, 65–70.

Kogan, N. (1969). Context effects in the assessment of creative ability in young children. In K. Ingenkamp (ed.), *Developments in Educational Testing*, vol. 1. London: University of London Press.

Kogan, N. & Morgan, F. T. (1969). Task and motivational influences on the assessment of creative and intellective ability in children. *Genet. Psychol. Monogr.* **80**, 91–127.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychol. Rev.* **69**, 220–232.

Pankove, E. & Kogan, N. (1968). Creative ability and risk taking in elementary school children. *J. Personality* **36**, 420–439.

Thurstone, T. G. (1963). *Primary Mental Abilities: Technical Report*. Chicago: Science Research Associates.

Wallach, M. A. & Kogan, N. (1965). *Modes of Thinking in Young Children: A Study of the Creativity–Intelligence Distinction*. New York: Holt, Rinehart & Winston.

Wallach, M. A. & Wing, C. W., Jr. (1969). *The Talented Student: A Validation of the Creativity–Intelligence Distinction*. New York: Holt, Rinehart & Winston.

Ward, J. (1967). An oblique factorization of Wallach and Kogan's 'creativity' correlations. *Br. J. educ. Psychol.* **37**, 380–382.

Ward, W. C. (1968). Creativity in young children. *Child Dev.* **39**, 737–754.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children: Manual*. New York: Psychological Corporation.

(*Manuscript received* 30 *January* 1970)

Cropley writes (30 March 1970):

Kogan points out that oblique rotation of the Cropley–Maslany data yields a 'pure' factor of creativity which has a positive but very small correlation with the two intelligence factors, the correlation of creativity with the first intelligence factor being of much the same order as the intercorrelation between the two intelligence factors (0·124 as against 0·286). Thus his reanalysis indicates that there is a distinct creativity factor which is slightly related to intelligence, the point which the original paper was intended to make. However, the magnitude of the inter-factor correlation derived from the Wallach–Kogan tests is strikingly small when compared with the creativity–intelligence correlation of 0·514, obtained through the use of oblique rotations, with earlier creativity tests (Cropley, A. J., 1966. Creativity and intelligence. *Br. J. educ. Psychol.* **36**, 259–266). The original treatment of the present data resulted in an unusually small general factor, and quite a close resemblance to a bifactor solution. However, rotations have been employed in other studies cited by Kogan, and have become a standard procedure, so that his use of oblique rotations, to obtain pure but slightly correlated factors, more forcefully emphasizes the greater independence of Wallach & Kogan's creativity tests from conventional intelligence tests than did reporting unrotated factors. For this reason, his reanalysis of the Cropley–Maslany data is helpful.

Br. J. Psychol. (1971), 62, 1, p. 119
Printed in Great Britain

PAPERS TO APPEAR IN FORTHCOMING ISSUES

(Not previously listed)

CROMER, R. F., M.R.C. Developmental Psychology Research Unit, Gordon Street, London. The development of the ability to decenter in time.

DI SCIPIO, W. J., Bronx State Hospital, New York, U.S.A. Divergent thinking: a complex function of interacting dimensions of extraversion-introversion and neuroticism-stability.

FRITH, C. D., Department of Psychology, University of London Institute of Psychiatry, Denmark Hill, London. Strategies in rotary pursuit tracking.

GALE, A., CHRISTIE, B. & PENFOLD, V., Department of Psychology, University of Exeter. Stimulus complexity and the occipital EEG.

JOYNSON, R. B., Department of Psychology, University of Nottingham. Michotte's experimental methods.

MARTIN, D. N. & WARDE, S., Assessment Unit, Red Bank School, Newton-le-Willows, Lancs. The performance of approved school boys on the Gibson Spiral Maze.

NICHOLSON, J. N. & GRAY, J. A., Institute of Experimental Psychology, University of Oxford. Behavioural contrast and peak shift in children.

POWER, R. P. & MACRAE, K. D., Department of Psychology, Queen's University, Belfast. Detectability of items in the Eysenck Personality Inventory.

SCHÖNE, H. & WADE, N. J., Abteilung Schneider, Max Planck Institut für Verhaltens-physiologie, Seewiesen-über-Starnberg (Obb.), W. Germany. The influence of force magnitude on the perception of body position. II. Effect of body posture.

SCHULMAN, A. I., Department of Psychology, University of Virginia, Charlottesville, U.S.A. Recognition memory for targets from a scanned word list.

SUTTON, A. J., School of Behavioural Sciences, Macquarie University, North Ryde, N.S.W., Australia. The use of quadratic discriminant analysis for the measurement of profile distance in social perception.

Br. J. Psychol. (1971), 62, 1, pp. 121–142
Printed in Great Britain

BOOK REVIEWS

The Bases of Social Behaviour: an Approach in terms of Order and Value. By PETER KELVIN. London: Holt, Rinehart & Winston. 1970. Pp. xvi + 334. 35*s*.

This is a book that faces the analysis of liveable lives from a fundamental level. Kelvin says that the basic need of men is to predict events in the world. Physical and biological reality provide us with fact-dependent constraints and laws of a high level of inevitability. But social phenomena are not bound by such laws and are potentially infinite in their variation. The analysis of their in-fact finite pattern is the subject matter of this book. These social patterns are value-dependent. Their man-made order is of some lower level of inevitability, and their subjective validity is based on some form of consensus. From the central concern for prediction come the order and consensus, and underpinning those, the values which integrate and justify them.

Kelvin acknowledges that the delineation of social orders falls properly to sociology and anthropology, and decisions of desirability to political and moral philosophy, and theology (indeed he is one of the few psychologists who even mention the place of history and religion in social material). But clearly the study of social behaviour, its development and forms, is social psychology.

What then does an approach in terms of order and value mean? Kelvin sees individuals as active beings, who are in some sense a function of the order which they perceive in their environment. It is in the analysis of order that Kelvin has made his most original contribution. Order in the simple, logical sense of meaningful categories is a *sine qua non* of perception and response. In our social environment, too, there is a series of orders of subjective probability which are built up from past experience. But, more importantly, Kelvin uses order in a second sense, that of an order of priorities. Intertwined with our probability order is a classification of objects, goals and paths in terms of importance, of rightness, and in terms of value to us. Humanistic writers on personality acknowledge that values provide the central core of the self. It is good to have a social psychologist make the point so forcefully.

In fact, the book then is a closely integrated argument that all the fundamental areas of social functioning may fruitfully be interpreted as contributing to order and are themselves ordered in terms of value. The argument is of a high level. Kelvin works cogently, honestly and painstakingly (occasionally obsessionally) at definition and conceptual analysis. The reader is taken through, for example, the meanings of compliance, stereotype or power; but also through what the phenomena 'mean' functionally to individuals and groups.

The continuity and power of this analysis is well demonstrated in the field of attitude and attitude change. Attitudes are said to provide the ultimate basis of social behaviour and so come first in the book. They are processes of ordering, in which we relate what we know and believe to what we do, modified by what we feel. In attitude measurement we sample basically two functions: intensity of feeling, and intensity of belief and behaviour. These two provide not only convenient ways for the scientist to categorize his data, but also two ways in which the subject organizes his world and his actions—an order of value and an order of probabilities. In so far as perception and belief and knowledge are themselves influenced by the desirability of possible events, the individual's order of value is seen as highly pervasive and primary. Consistency theories are discussed not in terms of their biological or god-given function, but as possibly instrumental aids for clearer predictability.

As can be imagined, the argument about order has particular application in the field of norms and conformity. It goes like this: for there to be effective interaction, we must have order; order is provided by norms; and norms are provided by probability expectations. Such shared guides may be formal, in which case they are associated with simple reward and punishment. But interest centres on the analysis of informal norms which are crucially associated with subjective probability orders. Behaviour connected with informal norms is prescribed because it is deemed to be valid. The validity is itself inferred from the frequency of occurrence of the behaviour in question. This circularity must be accepted as a fact. The frequency is a reflexion of a consensus that this is the proper behaviour. In the ordinary way, then, conformity occurs, not because of pressure to conform, but because conventional behaviour has consensual validity. This analysis

may be taken beyond words. Kelvin suggests provocatively that by this argument we have for the first time a means of measuring the *strength* of norms. Individuals may be asked to scale the probabilities of the occurrence of various kinds of behaviour—for the whole population, for certain membership or reference groups, or for themselves. Comparison among these will show how well an individual would 'fit' different groups, the state of his anticipatory socialization, and the degree of acknowledged divergence of certain groups.

In no area are particular details of research findings presented. For example, no experimental or 'casework' results of socialization studies are found in the socialization chapter. Rather there is a delineation of socialization as a process which makes the world predictable for *S*, which makes *S* predictable for others, and which develops and modifies our image of our self. In childhood this social learning mainly modifies behaviour; in adulthood it continues in terms of modification of attitudes and values, which are speculatively presented as generalized programmes which act as translation processes.

Good students will relish Kelvin's 'think piece'. The best heuristic application would be to recommend the book to be read once before and once after the usual course in social psychology. Harold Proshansky, in the foreword, suggests that it be married to a book of readings (he is too modest to cite one). But Kelvin himself provides an excellent commentary on further readings, with perspicacious notes (e.g. Goffman's style is easy and cynical, but his work is serious and important for the student).

The main thing to be said though is that this is an important contribution to social psychology. It makes us stand back from data, and even from theories, to look at what exactly we have to account for. As a reflective, but still empirical analysis, it seems a proper British contribution to a largely American discipline.

The publishers have made the book rather strikingly ugly to look at. The paper is grey. The binding makes it hard to open; once opened, the spine is broken. But it is cheap, even for limp covers.

HALLA BELOFF

General Psychology: Modeling Behavior and Experience. By W. N. DEMBER and JAMES J. JENKINS. Englewood Cliffs, N.J.: Prentice-Hall. Pp. xiv + 784. 100*s*.

Psychology. By B. VON HALLER GILMER. New York, Evanston and London: Harper & Row. Pp. xi + 512. 93*s*.

Introducing Psychology: an Experimental Approach. By D. S. WRIGHT, ANN TAYLOR, D. ROY DAVIES, W. SLUCKIN, S. G. M. LEE and J. T. REASON. Harmondsworth: Penguin Books. Pp. 736. 20*s*.

Psychology: a Story of a Search. By W. LAMBERT GARDINER. Belmont, Calif.: Brooks-Cole. Pp. x + 338. No price given.

Introductory texts in psychology often tend to be very much alike. This is very definitely not so in the case of two of the books listed above: more about each of them later. That leaves two others (those listed first and second) which conform more closely to the stereotype: large format, many illustrations, including some colour plates, and, no doubt partly in consequence, a price rather high even by present-day standards. There are, however, also points of difference, as we shall see.

Drs Dember and Jenkins open their introduction with the remark that 'one of the problems with a textbook is that there is no correct place to start. Everything that the authors want to talk about is better discussed after the reader knows everything else'. One can agree, and can add that even when one has chosen a starting-point, it is difficult, for similar reasons, to decide what comes next, and what, throughout, goes with what.

Setting aside Dr Gardiner's book for the moment, one finds that in spite of avowed differences in orientation and intention there is a fair amount of agreement between the various authors on matters of arrangement of at least the main divisions of their subject matter. The Dember and Jenkins' book is in five 'Parts', actually untitled but readily identifiable as dealing with, respectively: Psychology as a Science; Sensation and Perception; Learning; Cognitive Processes;

Motivation, Emotion and Group Processes. The four parts of Dr Gilmer's book are: Part One, 'Psychology and the Individual', including a section on personality; Part Two, 'The Biological Aspects of Behaviour', with sections on Bodily Mechanisms, the Human Senses, Perception, and Intelligence and Abilities; Part Three, 'The Dynamics of Behaviour', covering motivation, learned behaviour, emotional behaviour and cognitive processes; Part Four, 'The Social Aspects of Behaviour', with a section on Mental Health.

The part titles in the third-named book are rather different, although the order in which topics are presented is not unlike those just described. Following a relatively short introduction, Part Two, 'Structure and Behaviour', covers, among other things, the nervous system, the brain, and 'Needs and Drives'. Part Three, entitled 'The Use of Experience', encompasses perception, learning, skilled performance and remembering; Language and Thinking, often grouped with remembering as cognitive processes, receive separate treatment as 'Symbolic Behaviour' in Part Four. In Part Five, 'Individual Differences' covers intelligence, personality and 'normality and abnormality'. Part Six is devoted to 'Social Influences', and includes a chapter on Moral Development, a topic not separately treated in any of the other books.

It will be seen, then, that there seems to be a fairly uniform broadly conceived progression from the biological through the cognitive to the social; where there is less agreement is on the question of where certain topics such as personality and abilities fit in. More significant, or at least potentially so, are differences in intention and approach. The most explicit statement comes from Wright and Taylor, hereinafter 'Leicester', since the book represents the concerted effort, in a very true sense, of six members of staff at the Leicester Department of Psychology. Amplifying the book's subtitle, the authors admit to 'a comparative absence of theoretical discussion', having preferred 'to place the emphasis upon the analysis of concepts and the reporting of empirical findings'. In contrast, Dember and Jenkins lay very considerable stress on theory, or rather 'models'; 'modeling' in their subtitle would seem to mean *the construction of models or theoretical systems* of behaviour and experience. Nevertheless, they do not underrate the importance of reporting empirical data, and indeed their book has probably the widest coverage of the four in this respect. Gilmer's treatment may be seen as a compromise, but is clearly formulated as a 'systems approach', each of the sections into which the large 'Parts' are divided presenting research material and theory, followed by a discussion of 'problems relevant to daily living'. Psychology is presented as 'both a science and a profession', an aim not explicitly stated by any of the other authors—if indeed any of them had such a purpose in mind.

This thought raises many other questions, most of them having some bearing on your Review Editor's suggestion that 'the reviewer might care to reflect for a moment on the function of general texts'. Can an introductory text serve the interests of both the student and the general reader? If it is addressed primarily to the student, can it be useful beyond the first year (or, as in the case of Gardiner's book, beyond the single semester or 'quarter' for which it is said to be designed)? Can it stand on its own, and should it try to do so? What can be omitted? What is the function of illustrations? Is there any ideal format? And above all, perhaps, what should be the level of discourse?

This last question in a sense subsumes most of the others (which I have not attempted to answer systematically). I have posed it separately because it gives an opportunity for introducing Gardiner's book, on which I have as yet made no comment. The book is printed on yellow paper, because 'psychologists have demonstrated that black on yellow is the most readable combination', with extra wide margins so that the reader can add comments and thus 'talk back' to the book. The text consists of the author's 'introductory lectures in psychology written down'. Asides (i.e. jokes, some of them chestnuts indeed) are included, though relegated to footnotes; the style in the main text too is marred by an unrelieved facetiousness which in the context of a serious book is so inconsistent with one's expectations as to make the book almost unreadable. There is neither a table of contents nor a glossary, the former being replaced by a diagram of the 'hierarchical structure' of the book which determines the rather idiosyncratic order of presentation, and the latter by a 'chart' clarifying the technical terms used in the book, for which some special virtue appears to be claimed. It would seem that the author has fallen over backwards in his efforts to achieve novelty and avoid dullness so completely that, to extend the metaphor, he has been unable to get up again.

Gilmer's writing also occasionally verges on the colloquial, or rather, conversational, while Dember and Jenkins' text is probably about average for its kind. It is a pleasure to note the clarity of exposition in the Leicester book, which amply demonstrates that serious writing does

not have to be jargon-ridden in order to be scientific, nor colloquial in order to be readable. It is gratifying, too, to find how much use has been made of British sources: I shall leave the reader to play for himself the game of searching for names of cisatlantic psychologists in the indexes of the other books under review.

The reader will have gathered that the Leicester book is heading for nomination as 'Best Buy' in the *Which?* sense. Quite apart from its positive merits, its price even for a paperback is amazingly low. This has no doubt been made possible partly through the sparing use of illustrations, these being confined to line drawings and diagrams: all are models of clarity. The two large-format books have, as we might expect, a greater profusion of illustrations. Those in Dember and Jenkins are outstandingly well produced, and include nothing superfluous. Not quite the same can be said of Gilmer: a representation of the 'colour circle' seems to include at least one clear violet among the non-spectral hues; and is there really anything to be gained by printing, for example, a half-page photograph to illustrate the remark 'a classroom climate can be stimulating to some and bring on boredom in others'? In Gardiner's book the main illustrations are very small reproductions of 'custom cartoons by Tony Hall' which aim to 'make pungent comments on the ideas'.

In his epilogue Dr Gardiner calls attention to certain topics (notably personality and applied psychology generally) which are not covered in his book. Also missing is any detailed discussion of sensory phenomena, an omission one notes also in the Leicester book. If this subject matter is considered essential in a first-year or one-year course, then the other two books are 'complete' in a sense in which Leicester and Gardiner are not. Neither Dember and Jenkins nor Gilmer, however, would claim what one might call 'self-sufficiency' for their books. Each has an associated work book or study guide (which was not available to the reviewer). Furthermore, while discussing problems of comprehensiveness versus selectivity, Dember and Jenkins refer to 'the level-headed instructor in whose course the book is being read'. Such a person, or a tutor who is genuinely involved, is, in this reviewer's opinion, an essential adjunct to any introductory textbook of psychology. I am not, of course suggesting that one book is as good as another. This review, it is hoped, will perhaps be helpful in making a choice among the four books listed.

BORIS SEMEONOFF

Attention: Selective Processes in Vision and Hearing. By NEVILLE MORAY. London: Hutchinson. 1970. Pp. xiv + 218. 40*s.*; paperback, 20*s.*

'Attention is back again', declares Moray, with the enthusiasm of a fashion editor announcing the autumn collection of granny dresses and lace-up boots, 'respectable both in theory and in practice'. Attention is, indeed, an 'in' subject, given fresh impetus in the 1950s by the work of Colin Cherry and Donald Broadbent and sustained in part by the availability of stereophonic tape-recorders, in part by an excellent series of Dutch symposia on attention and performance (the first was in 1966, not 1967 as stated in the text) and the enthusiasm of a small group of research workers amongst whom Neville Moray has been one of the most productive. This book takes Broadbent's *Perception and Communication* as its starting point and reviews the literature up to date (including a number of stop-press footnotes). Even given the relatively recent publication of Neisser's stimulating *Cognitive Psychology* the need for a text drawing together work during the last 12 years is apparent.

With brief reference to the history of the concept of attention the author characterizes the field of study as to do with a set of situations and phenomena such as mental concentration, vigilance, selection, search, and set, all bearing no more than family resemblances to each other. Attention is not a unitary problem and in this sense a comprehensive theory of attention is not expected to emerge. In fact the experimental and analytical effort described in the book is concentrated almost exclusively on the problems of selectivity and switching with listening experiments very much to the fore.

As an aid to analysis a notation system for describing experiments is introduced which the author himself admits is somewhat cumbersome. This is virtually abandoned later in the book but nevertheless does make explicit the variables which are involved and the omission of critical controls in some experiments. It is significant that a set of recent experiments described in the last chapter involve simple precisely specified signals, rather than verbal messages, and detection rather than repetition responses. The pursuit of rigour often seems to lead in this direction.

Six current theories of selective attention are outlined, all too briefly, and the bulk of the book contains reviews of auditory and visual selection and the problem of switching attention. After all, Moray proposes a modified version of Broadbent's original theory based on discontinuous sequential sampling of variable duration. Selectivity is regarded as operating not on relatively raw inputs but on representations which are already coded into some internal language of the nervous system which permits cross-correlation between 'similar' or 'related' inputs. Whilst Moray's model has some similarities with Broadbent's original hypothesis, some things do seem to have changed considerably over the years. For instance, the idea of a limited channel somewhere in the centre of the system is now rather different. A limitation there is, but only in 'overall processing capacity'. The other development evident, not only in Moray's proposal but in Treisman's and Neisser's work, is in the increasingly important role assigned to pattern recognition in the attentional process. Indeed, as I read the book I began to wonder if the whole notion of selection, so long conceived as a switching operation somewhere in the system, is not simply the inevitable result of a hierarchy of pattern recognition processes. For example, an object may be recognized as a cube, or as a red cube or as a red wooden cube. If it is recognized as a cube (say, for example, one is searching for cubes) there is a sense in which its redness or its woodenness are not attended to. Is it possible we may be mistaken in writing pattern recognition and selection into separate boxes in our imaginary flow diagrams?

Although a theory is presented Moray allows his honesty to temper his enthusiasm. Attention is a many-faceted concept and it is not even very easy to reconcile the data on visual attention with those in selective listening. But the final blow, reserved for the last page, is that the practice variable has been almost totally disregarded in research on which these suggestions are based and yet it is known that subjects can achieve high levels of skill in these experiments. Any theory at this stage must remain only a theory about the behaviour of experimentally naive subjects. The series editor, Kevin Connolly, characterizes the aim of the Hutchinson monographs not as 'definite (*sic*) texts but rather a state of the art analysis' and this is certainly true of the first example. This is a useful and stimulating book about work still in progress and ideas which are still developing.

JOHN ANNETT

Diseases of Attention and Perception. By MONTE JAY MELDMAN. Oxford: Pergamon Press. 1970. Pp. xi+241. 100*s*.

Pathology of Attention. By ANDREW MCGHIE. Harmondsworth: Penguin Books. 1969. Pp. 192. 6*s*.

Here are two books which, although superficially very similar, differ markedly both in content and expressed aims. Meldman attempts an ambitious wide-ranging application of the concept of attention to mental (as well as physical) illness, operating largely on the psychodynamic level. McGhie, on the other hand, offers a more modest review of the advantages of human performance methodology in clinical diagnosis.

Diseases of Attention and Perception fails, largely, I think, through its use of the concept of attention, which is far too broad to have any real explanatory value. '...the brain...performs one function...attention' (footnote, p. 177). The mode of operation of this function is difficult to deduce from the book, but there seems to be a fixed amount of attentional 'energy' which is distributed between all the bodily functions (walking, breathing, digestion, perspiration, as well as cognitive processes) in characteristic ways in different diagnostic groups. This, as I understand it, is the bones of the theory. In the same footnote Meldman adds that 'Most readers will either not understand this concept of attention or they will disagree with it, for a traditional use of the word attention refers only to selective attention as a conscious and more or less voluntary event. Use of the term [here] is much broader...and to some may seem an unjustifiable oversimplification. In the last analysis, what matters is whether [it] is useful in developing more rational forms of psychiatric diagnosis and treatment'. I must include myself in the group of readers referred to, and I have very strong doubts about the eventual usefulness of the approach, though I am no competent judge of that.

As to the actual content and organization of the book, it is too fragmentary, with headings and subheadings on practically every page. There are ten chapters, though only in three or four of them is there very much clinical psychology. There is an abundance of physiologizing, and

whole tracts of pure physiology with little direct relevance to psychological illness (for example, no less than five schematic drawings showing the anatomical organization of the reticular and limbic systems in the chapter on 'The Form and Functions of Attention'). It is easily arguable, of course, particularly since this is an early chapter in the book (chapter 3) that such detail is necessary groundwork for the development of the theory in later chapters. This is just not the case, however. It is simply a chapter on physiology/anatomy which happens to have been dropped into the book at that place. Further chapters (notably 4 and 6) carry on the tendency to include largely unnecessary material of this kind, so that a clinical psychologist interested in *behavioural* malfunctioning may easily be excused for loss of patience. One can have no quarrel with attempts to provide physiological explanations for even such eminently behavioural problems as those of mental illness, but there must be some attempt to integrate the two sets of data in a meaningful way. The inescapable conclusion is that the author, feeling insecure about the acceptance of his theory, finds a need to dress it in respectable scientific clothes. Hidden amongst the physiology there is, however, reasonable coverage of the literature on the differences in cognitive and perceptual behaviour between normal and clinical groups, though it is significant that only about a third of the book is given over to a description of these differences.

The last two chapters describe in more detail the author's original contribution to diagnosis, the 'attentional profile'. In this, disease is regarded as an imbalance between the level or amount of attention allocated to the various subsystems of behaviour; the profile is a visual representation of this pattern of attention, and is the primary diagnostic tool. Thus, anxiety reaction is typified by an excess of alertness (synonymous here with arousal), more than normal attention to emotional and visceral aspects of behaviour, and less than the fair share of selective attention! Nineteen such profiles are provided at the end of the book, of which about half a dozen seem to be actually distinguishable (the only discernible difference between the anxiety reaction and the phobic reaction is that the latter invests three times as much attention in 'action').

I feel that there must be something in this book which I have missed. It seems to promise so much at first glance, but has little coherence or substance on further exploration. I looked for consolation to the index, but this was worse. There is no separate author index, though there are 22 references to authors intermingled with the subject items in the index. (There are about 400 authors mentioned in the bibliography!) Admittedly, some of this elite 22 are those who give their names to headings in the text (Deutsch's theory, Hebb's model, etc.), but most seem to be drawn randomly from a hat. Zung is given in the index as p. 237, which is his bibliography entry (whether it is his *only* entry is difficult to say). Freud, who crops up regularly throughout, is given as pp. xiii, 23, 24, 27. There is no page xiii, though Freud is indeed mentioned in the text on pp. 23 and 24, but not on p. 27. In addition he crops up on, e.g., pp. 25, 26, 49, 68. On p. 13, Petrie, Silverman and Sullivan are cited, though only Silverman appears in the index. It is easy to go on, but it would be pointless.

It is with relief, therefore, that I turn to Andrew McGhie's book, *Pathology of Attention*. It is not possible to say that it is a better book, since it is vastly different, but it succeeds in a modest aim where Meldman's fails in an ambitious one; it is another example of the good quality 'Science of Behaviour Series'. In six chapters, McGhie discusses models of attention, then, in turn, experimental studies of attention in neuroses and psychopathy, affective psychoses, schizophrenia, brain damage and mental deficiency. The presentation is clear and the discussion of findings particularly good. This kind of approach to clinical diagnosis, the identification of systematic patterns of behavioural impairment, is undeniably useful. Clinical psychologists have for too long been limited to administering routine tests without thinking about the psychological functions which are involved. With the sophisticated test methodology available today there is little excuse for not asking more specific questions like 'Is short-term memory impaired, or long-term memory—or both?' or, not simply 'Is attention impaired?' but 'Is the impairment one of selectivity, or expectancy, or one which only appears after some time at the test (vigilance)?' This book, although not going a long way along this path, will be useful in acquainting clinicians with such techniques, as well as encouraging those already using them.

To summarize the relative merits of these two books, Meldman's, I feel, is a less useful contribution to experimental psychopathology, it costs nearly 17 times as much and contains a far inferior index, though it does have the more attractive cover design.

G. R. J. HOCKEY

Attention in Neurophysiology. Edited by C. R. Evans and T. B. Mulholland. London: Butterworths. 1970. Pp. xxiii + 447. £10.

The papers in this book were presented at an international symposium held at the National Physical Laboratory in 1967. They provide an important review of techniques and of the results of attempts to link the behavioural phenomena of attention with underlying neurophysiological mechanisms.

The tremendous increase in interest in attention in the last few years among behavioural psychologists has in the end led us to realize that attention is not the name of a unitary concept, but that many and varied mechanisms are involved in the phenomena which we subsume under that label. Until recently physiological research had not contributed a great deal to the renaissance of attention, with the exception of the well-known experiments of Hernández-Peón and his group, which have subsequently been discovered to have been badly corrupted with artifacts. Perhaps the most developed area of physiology had been the work on the reticular system. But this symposium makes one feel quite optimistic about the future.

Summarizing some of the most important opinions which seem to have been generally held, we may say that there is little or no evidence for the peripheral gating of information which Hernández-Peón thought he had observed, unless the special case of the suppression of information intake during saccadic eye movements is included. Alpha rhythm, so long a prominent contender in the attention stakes, can now be firmly ruled out as of any direct interest. Overall, one gets the picture that the techniques now available are perhaps able to match the sophistication of the mechanisms of the phenomena being studied.

Of particular interest is the increasing evidence about the nature of contingent negative variation (CNV) which was originally called the 'expectancy wave'. This seems to bear a real, if complex, relation to attentional states, and perhaps to set. The work is clearly of great importance.

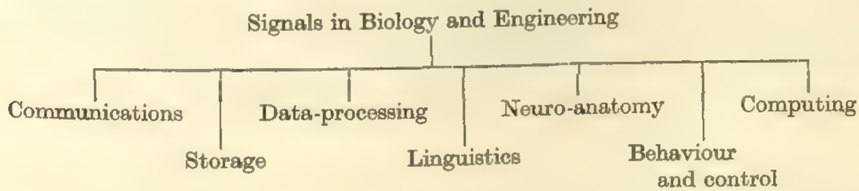
The symposium, whose guest of honour was Lord Adrian, provides fascinating and important reading for those interested in attention. It is unfortunate that the publishers have priced the book out of the pocket of most readers. Even to libraries the price seems outrageous.

NEVILLE MORAY

Encyclopaedia of Linguistics, Information and Control. Edited by A. R. Meetham and R. A. Hudson. Oxford: Pergamon Press. Pp. xiv + 718. £15.

There is probably an optimum degree of disorganization for a reference book. If the exercise of editorial functions is too rigidly systematic, valuable information may be excluded, especially about peripheral topics or about relations between topics, while if it is insufficiently systematic, difficulties of labelling arise and wastage will occur because information is inaccessible or duplicated.

It is quickly apparent that the editors of this encyclopaedia have not worked to too rigid a system. The foreword says: 'To expose the common feature of all the articles it is only necessary to point out that the signal is as fundamental in its own way as the atom'. There follows a diagram adapted below, with the suggestion that this is how the encyclopaedia might be classified.



One can see that this scheme does not tie editors or contributors down too tightly and allows plenty of opportunity for happy accidents to occur. Is it then too loose a scheme for the information in the encyclopaedia to be satisfactorily organized?

As far as the subjects listed in the title are concerned, the scheme does in large measure succeed, the coverage being comprehensive and the information accessible.

Under 'Control' there are many articles that can be recommended, including those on adaptive systems, on homeostasis, on various forms of control (predictive, on–off, sampled-data) and others besides. There is one lack, and that is of a good introduction to the subject. The article, 'Control: Basic Element', perhaps intended to meet this need, does not do so. It is written from the point of view of process control, starts from a particular example without making it clear that it is not a general case, introduces the Laplace transform without reference to classical methods, gives an insufficient definition of linearity and includes no references.

Under 'Information' (including computing) there are again a large number of useful articles: on recognition, simulation, automata, unsolvable problems, computer-operating systems and on different kinds of computers, e.g. stochastic. To be noted is an approachable account of pseudo-randomness, in Hartley's article on traffic simulation. And this time there are good introductory articles, by Woodward, on communication theory, and with Wetherall, on random signals and noise.

These make one regret the absence of an article by Woodward on detection theory, since his work with Davies demonstrating the essential part played by posterior probability in detection is the largely unacknowledged starting point of the revolution in psychophysics that has taken place in the last 16 or 17 years.

Unfortunately, Woodward's lucid exposition did not reach psychology in its original form, but in a reworking by Birdsall and Peterson of such deadly rigour as to be all but unreadable: so that psychologists have in general had to make use of an oversimplified account of the detection process, giving the logarithm of the likelihood ratio as the decision variable without qualification, and making no mention of posterior probability. We are still suffering the consequences of this historical accident. One of them is an article in this book by Treisman, on sensory discrimination, which is an extended exercise in bathwater-retrieval.

A large number of helpful articles can be found under the label 'Linguistics': on bilingualism, phonology, phonetics and on the related topics of speech perception and analysis. A characteristically all-embracing article by Good on the statistics of language includes, if not the kitchen sink, Pareto's law of incomes! One must, however, note important omissions, concerning the development of speech in the child, and the pathologies of speech—nothing, for instance, on brain damage; so that 'aphasia' does not appear in the glossary/index.

These omissions turn out to be symptomatic. One begins to see that as far as psychological, and to some extent, physiological topics are concerned the editors' classification scheme is not sufficiently systematic. Behaviour is not something to be set alongside communication, data-processing, storing and computing, as they have done. It *depends* on those activities and will not occur in their absence. The editors have failed to appreciate that one cannot write effectively about behaviour without thinking about how it is organized. The result is a somewhat haphazard selection of psychological topics and inadequate editorial control of contributions. For instance, the article 'Perception, visual' by Crawford, is about almost every aspect of vision except perception; the article on attention is quite unsatisfactory; two articles by Whitfield contain no references—indeed, one on cochlear mechanics manages to avoid mentioning von Békésy; stabilized retinal images crop up in four articles, the Ganzfeld in none; and the glossary/index includes 'psychological refractory period', 'Pulfrich's phenomenon' and 'range effect' but not 'conditioning', 'visual constancy' or 'vigilance', to give just a few examples.

It probably takes an encyclopaedia several editions to settle down. This one contains more than its share of errors. Gaussian variable, normal ogive and Poisson distribution are all incorrectly defined. In addition to these errors of content, there are many others that are due to slack proof-reading. Nevertheless, I am sure that a wide variety of users will find, despite moments of exasperation, much that is interesting, unfamiliar and useful to them in this book.

I. M. HUGHE

Cybernetic Machines. By T. N. NEMES. London: Iliffe. 1969. Pp. 260. 75s

This book, edited in its English edition by W. A. Ainsworth, is superbly translated from an original Hungarian work. Apart from being very readable and well illustrated, it is welcome as an academic contribution to the literature of cybernetics in closing a number of gaps in the recent history of the subject. Nemes died in 1960 at the age of 65, prior to the publication of the two Hungarian works which form the basis of this book and taken together, the circumstances of his life span, engineering training and nationality, combine to give the work a rare value.

Although much of the material is familiar, that which is not has its roots in the cultural preoccupations of Eastern Europe during a period which stopped short of the boom in electronic computing. Hence the vivid descriptions of automata for musical composition, chess playing and demonstrating genetic mechanisms are from a man who was a contemporary and compatriot of many of the inventors and to whom some of these devices represented the ultimate in mechanization. Some of his phrases have a strange ring to Western ears, reflecting the era and climate of opinion in which he lived and worked; for instance, 'Logic is not to be considered a part of dialectic materialism, the philosophy of Marxism Leninism;...dialectic logic also investigates thinking and reasoning but from a different point of view'.

Although cybernetics is a predominantly theoretical subject, it, like theoretical physics, has drawn much of its inspiration from both natural and constructed models. Indeed it is arguable that tangible models contribute more to the prevailing climate of opinion in a subject than the ponderous rigour of theoretical proof. It is not merely that hardware exemplifies and so advertises abstractions but that mechanisms demand discourse relating to their manufacture and in some cases to their use and maintenance. By this means their features become established in terms of language, whose frequent use conditions our acceptance of conclusions which in the abstract cannot be brought into conceptual focus. Though it is sometimes argued, for instance, that computers owe their ancestry to the cumbersome theoretical automata of Turing and his contemporaries, yet an inspection of their lineage argues a much humbler descent via electronic technology and the engineering 'Gestalt' which accompanies switching circuits used in communications systems.

This theory hardware dichotomy is reflected in the composition of Nemes' book, for while the first third is devoted to a treatise on formal logic the latter section deals exclusively with hardware of the type which laid the foundations of cybernetic philosophy. Predictably the logic fails to do more than describe the obvious in concise terms and fails to connect with the rest of the book. Significantly the philosophical discussion of implications accompanies the hardware rather than the theory.

This unusual juxtaposition is directly attributable to the composition of this edition from two separate source works, but though these were published within the same quinquennium the shift in register is so marked that one feels they must represent different periods of the author's own intellectual development. It is almost as though Nemes, with his strong engineering background, sensed the abortive nature of formal logic and decided to reset his philosophy in the more versatile framework of hardware and conceptual models. Be that as it may, the main development occurs in the second section, where Nemes' descriptions of machines and algorithms are presented with an authority and respect for detail which set a new high level in this field. Surprisingly, a high proportion of the references are to devices of Western origin and even here we find an accuracy which is rarely surpassed by Western commentators.

The only adverse criticism of this book is that, in attempting to cover such a wide range of mechanisms so concisely, Nemes has broached a whole new series of questions on which one would like to hear his discussion. Perhaps future translations may throw some more light on his work, but in the meantime this book forms a very fitting, though perhaps unintentional, memorial to a man who deserves much wider recognition in Western circles. BERNARD CHAPMAN

Developments in Applied Psycholinguistics Research. Edited by SHELDON ROSENBERG and JAMES H. KOPLIN. New York and London: Collier-Macmillan. 1968. Pp. vii + 311. 84*s*.

Most students of psychology are well informed about the traditional studies in verbal learning and have a rudimentary acquaintance with the arcana of contemporary psycholinguistics. Less well known are the applications of this knowledge in the realms of first- and second-language teaching and the language impairments associated with deafness, aphasia, schizophrenia and mental retardation. This volume is therefore to be welcomed as a useful and, in parts, stimulating introduction to such topics. Its link, however, with what is currently understood by 'psycholinguistics' is rather tenuous.

Following an able and succinct introduction by Koplin, the first of two chapters on 'normal' language topics is a review by Griffin of the development of syntactic control. Griffin's interests are primarily pedagogical and so his attention is focused on the language development of older children. Much of this work may be unfamiliar, even to those who rub shoulders with Brown, Bellugi, Ervin-Tripp and Slobin. It is, however, of considerable interest and deserves a wider audience. The chapter by Lane on second-language learning is largely an account of his own work on the use of operant procedures in training discrimination and differentiation of phonological features. These appear to hold some promise but one feels that Lane too frequently and optimistically invokes the concept of transfer when lacking empirical evidence. As he himself admits, the optimism is not always justified.

Four chapters on language pathology follow. Blanton, concerned with deafness, reviews at length the relations of language and thought. Alas, he finds the problem no more tractable than have the legions who have gone before him. What he does succeed in doing is to show the subtle links between evaluative language and the socialization of the deaf. Goodglass's chapter on aphasia is notable for its account of the significance of prosodic features in aphasic impairment; notable, too, for the suggestion that the sentence, 'The soldier write home every week', is unambiguously incorrect. Cromwell and Dokecki contribute an excellent review of the symptomatic and taxonomic problems of schizophrenia. Their linguistic research relies almost exclusively on simple verbal learning tasks, on the basis of which they present an interpretation of schizophrenia as being characterized by difficulties in disengaging attention. This is not altogether an original idea but in the hands of Cromwell and Dokecki is convincingly argued. Spradlin offers a cautious but disappointing account of the operant control of simple verbal responses in retarded children and the threads of the book are neatly woven together by Rosenberg in his overview.

Rosenberg and Koplin have done a good job as editors, but several proof-reading and typographical errors survive: Figs. 3–3 and 3–4 are wrongly numbered; on p. 159, 'twenty-four' should read 'twenty-five'; at least one of the probabilities in Fig. 5–2 is wrong; on p. 245, 'complimentary' should read 'complementary'.

This book would be a valuable addition to anybody's psycholinguistics library. The only general criticism is its almost total neglect of crucial problems in semantics. BRYAN BETT

Psycholinguistics. By JAMES DEESE. Boston: Allyn & Bacon. 1970. Pp. x + 149. 26*s*.

The aim of this textbook is to provide a brief beginner's introduction to current research in psycholinguistics. In many ways, it does this well. It is clearly and simply written and carefully explains all new terminology as it is introduced. The overall approach is thoroughly modern and a good range of material is covered within a small book. The weakness of the book lies in its tendency to superficiality and oversimplification.

The author, long famous for his studies of verbal behaviour from a neo-associationist viewpoint, has become a zealous convert to the new theology deriving from the linguistic theories of Noam Chomsky. This commitment is manifest from the very first chapter. Often the author's adherence to Chomsky's doctrines is completely uncritical. For example, we are told (p. 11): 'A child has as part of his native equipment a device embodying a linguistic theory of a high degree of complexity', and, further down the same page, 'something much more abstract and

difficult than language itself is innate. It is something that is best described as a universal grammar or as a device for producing language'. These statements are most unhelpful for several reasons. The author offers no definite evidence in support of these very strong claims. On the contrary, he presents them as if they were self-evident truisms, accepted by all. This is seriously misleading, since much of the most heated controversy, and some of the best research, in psycholinguistics has concerned exactly these hypotheses. Far from representing fundamental and well-understood facts, they are the source of many very open questions. Chomsky has claimed that all languages can be described by grammars sharing certain similarities. Even if this claim is upheld, it does not solve any psychological problems, but poses a host of difficult new ones. The psycholinguist must try to discover what psychological mechanisms cause all languages to be sufficiently similar as to be describable by grammars which, though differing in many respects, nevertheless have certain common features. To utter vague statements about children possessing innate grammars, and to equate these, as Deese seems to do, with a description of the working of the cognitive processes by which the child acquires its native language, is to confuse the true nature of the problem. All too often, Deese swallows the ideas of Chomsky and his followers whole and fails to digest them.

The chapter dealing with meaning is probably the best in the book. It is also the one which owes least direct debt to the ideas of generative grammar. Here the author provides rapid outline sketches of the componential analysis of meaning, the theory of semantic categories and the semantic differential. It is in dealing with these topics that he shows most originality.

If you want a brief look at what psycholinguistics is about, then read this book. If, however, you wish to acquire a thorough understanding of the subject, then it is still necessary to go to the specialized works listed in the bibliography or, perhaps best of all, to the beautifully written chapter by Miller & McNeil in *The Handbook of Social Psychology*. We still await a really good up-to-date general textbook on this subject.

R. Q. GOODWIN

The Acquisition of Syntax in Children from 5 to 10. By CAROL CHOMSKY. Cambridge, Mass. and London: M.I.T. Press. 1970. Pp. viii+126. 56s.

This book presents the published version of Mrs Chomsky's doctoral thesis, written at Harvard under Roman Jakobson. It is a most interesting piece of work and will surely be essential reading for all those working in the field of developmental psycholinguistics.

It had been widely assumed that, by the age of about six years, children had almost completed learning the syntax of their native language. In the light of this belief, most research was concentrated on the preschool child. Mrs Chomsky has looked beyond this age range and found clear evidence that in fact some of the more complex syntactic structures are not fully mastered until up to 10 years of age. Her thesis is an experimental exploration of comprehension. In this respect also, it goes beyond previous research, which had been almost wholly observational. The experimental procedure consisted of interviews in which the children were put through a routine of questions and commands relating to other children and to various dolls on the table before them. Thus the methodology shows clear similarities to that of Piaget. The data are presented very fully, with lengthy reproductions of the actual interviews. Again as in the work of Piaget, they are subjected to a thorough qualitative as well as quantitative analysis.

Four areas are studied in which the child's grammar differs from that of an adult—the construction of complements for the verbs 'promise' and 'tell' and for the verb 'ask', pronominalization, and the oft-cited distinction between 'easy to please' and 'eager to please'. For each of these cases the author is able to trace a clear succession of stages through which the child's comprehension develops. Despite wide individual differences in the ages at which these stages are attained, their chronological order of attainment appears not to vary. She also tries to draw conclusions from her data concerning the form of the processes underlying syntactic development. Clearly there are many fascinating problems here and we shall certainly be seeing a great deal more research along these lines in the future.

The book is well produced and contains charming illustrations of the experimental materials. Unfortunately, the price is very high for such a slim volume, even when it is as valuable as this one.

R. Q. GOODWIN

Verbal Conditioning and Behaviour. By J. P. Das. Oxford: Pergamon Press. 1970. Pp. vii + 163. 52*s*.

The claim on the cover of this book is that 'it clarifies certain key issues in the area of verbal behaviour focusing both on problems which are continuous and discontinuous with animal learning research'. The general strategy taken by the author is to present a brief discussion of a number of issues with his own research findings located in the appropriate chapters. It has produced a rather limited and unsatisfactory book.

In the first chapter, the author pursues the view that verbal behaviour can be treated as conditioned responses; difficulties in the way of this assumption are ignored. This is followed by a limited look at acquisition and extinction in verbal conditioning in which the problem of awareness as a factor producing changes in performance appears; its relevance to certain situations is agreed but it is argued that performance increments can occur in the absence of 'awareness'. The evidence and argument used to support this view are very weak, and the implications of 'awareness' for the type of explanatory model used is ignored. Next follow chapters on reinforcement and performance and probability learning. In the first, motor performance in a discrimination task is studied. This experiment attempts to repeat findings of Ramond (1956) with rats that the probability of choosing a stimulus which is reinforced twice as frequently as the other, first increases and then decreases. Success is claimed although slight modifications of the reinforcement ratio got rid of the effect. A section on social reinforcers completes the chapter; social deprivation is posited to be a real need state and the larger effect of 'good' than 'bad' on response latency for institutionalized as compared to normal children is used as support for this. The possibility that normals were responding so quickly in both cases that little difference could appear, is mentioned and appears to be a more economical way of evaluating the results but fails to find favour with Das. The similarity of probability learning to classical conditioning is then pursued, with the subjects' lack of awareness of the bases of their matching behaviour emphasized.

With chapter 5 comes a change of direction. It begins with an analysis of referential meaning in which Noble's 'm' value and the semantic differential (S.D.) score are evaluated. Not surprisingly, they are both found to be wanting, but the introduction of the S.D. allows Das to consider affective meaning and its persistence even under changed conditions from those in which acquisition occurred. The consideration of the acquisition of referential meaning as a conditioned response follows; some objections to this view are briefly mentioned but not pursued, and the demonstration of the validity of this classification is attempted by showing experimentally that because meaning is a conditioned response, transfer and generalization will take place. The experiments are concerned with the transfer of evaluative labels (e.g. good/bad) from one set of nonsense syllables to others, and to meaningful items. The theoretical strategy and the empirical results do not convince. Semantic satiation is next considered. Beginning with satiation as measured by changes in the S.D. score, the discussion quickly moves from a consideration of the 'cortical inhibition' hypothesis to one introducing attentional factors, with the cessation of a general orienting response implicated in the satiation phenomena. Empirical differentiation between these theoretical views is admitted to be very difficult. Das's solution is to proceed by assuming the near synonymity of semantic satiation and attention loss and then to measure satiation in terms of attention decrement. Such a method is claimed to establish the validity of the underlying hypothesis. The problems in this interpretation are ignored.

This is followed by a consideration of hypnosis as a form of conditioned response (although hypnosis is more than a conditioned response) with a susceptibility to be hypnotized and conditioned considered as stable personality characteristics. It is unclear whether they are thought of as separate or overlapping, but hypnosis is claimed to be a 'learned state of partial cortical inhibition', which suggests the latter. Whichever the choice, it does not seem to be a particularly helpful redescription. The chapter ends with a discussion of a possible relationship between hypnotizability and suggestibility. Highly suggestible and hypnotic subjects are also assumed to be weak on semantic satiation. If you satiate on the instructions you can hardly be affected by them! The penultimate chapter is a general plea for a look inside the black box, with among others an unconvincing example of the use of EEG measures as support for Pavlovian excitatory and inhibitory concepts. The last chapter looks to the future.

This book is difficult to read but this is not the main reason for its unsatisfactory nature. Das

argues that scientific advance comes only with an appropriate theoretical orientation within which the 'pointer' readings of the empiricist can be interpreted, and that his is conditioning. Conditioning is so variously and imprecisely described, however, and it here includes Pavlovian contiguity conditioning, instrumental conditioning and instructed voluntary conditioning, so that this helps not at all. The type of model implied by these labels is clearly varied but not defined, and this is typical of the conceptually confused way in which Das pursues his varied themes. In no case is there a sufficiently careful, detailed or extensive treatment to convince the reader that all the empirically relevant material and the necessary theoretical sophistication have been utilized. The uncritical use of many concepts, for example reactive inhibition, inhibition (undefined) and cortical inhibition adds to the confusion. It is clear, too, from the summary that the book is rather esoteric in the choice of topics included under the verbal conditioning and behaviour label; I have argued its other inadequacies and it cannot be recommended.

W. A. MATTHEWS

Early Learning in Man and Animal. By W. SLUCKIN. London: Allen & Unwin. 1970. Pp. 123. 35*s.*

The notion that personality is a function of early learning experience has not, for the human species at least, proved as easily demonstrable as was thought some decades ago. The many Freud-inspired investigations of early socializing practices were unable to find clear-cut effects in later life, and it is therefore hardly surprising that for a time scepticism was raised as to the possibility of being able to influence the behaviour of very young organisms at all. A direct examination of the learning capacities of infants was thus called for, and a voluminous literature on this topic has now come into existence.

In this book the various approaches to early learning are brought together. As the author points out right at the beginning, the field is a vague and imprecisely defined one, and indeed both 'early' and 'learning' are difficult to specify. Sluckin has, however, chosen to cast his net widely and included a considerable range of topics that, in one way or another, may be associated with infantile experience—topics such as imprinting and sensitive periods, classical and instrumental conditioning in infancy, gentling and shocking, imitation and the beginnings of language. The fact that this is a slim volume is no reflexion of the amount of data to which it refers; Sluckin has purposely treated his material at a superficial level in order to give students and other general readers a bird's eye view of the subject, and as such the book is a most useful contribution. It picks out and summarizes the various themes that have emerged in each area, points to gaps in our knowledge, and builds bridges wherever possible. In particular, it emphasizes the usefulness of a comparative approach and, moreover, it stresses the role of perceptual learning in the earliest stages of ontogenetic development.

Can we conclude that early learning does take place and that it does have a special character? No unequivocal answer can yet be given to these questions. It depends on one's definition of the terms involved, on the procedures employed, on the measures taken. For example, in 1931 Marquis wrote an article entitled 'Can conditioned responses be established in the newborn infant?' At that time the answer was thought to be a matter of a straightforward yes or no; since then, however, it has emerged that some response systems are more sensitive than others, that operant conditioning is more easily brought about in the infant than classical conditioning, that much depends on the nature of the reinforcer—and so on. Given optimal conditions, certain modifications of behaviour can already be demonstrated in the neonate; whether these are properly regarded as learning becomes a matter of semantics. What we do not know is what part such modifications play in the life of the developing organism, whether they have any significance outside the laboratory setting, and whether they are anything but purely temporary phenomena of no relevance to the formation of personality. No wonder that almost every section in this book concludes with a plea for more research—annoying, no doubt, to the student who wants quick, clear-cut answers to all problems, but in fact an accurate comment on the state of affairs in this area.

H. R. SCHAFFER

Direct Observation and Measurement of Behavior. By S. J. HUTT and CORINNE HUTT. Springfield, Ill.: Thomas. 1970. Pp. xii + 224. 138*s*. 6*d*.

In the last six years the authors and their associates have published many interesting reports on systematic observation of human behaviour among various classes of subjects and in various types of situations, which include exploratory behaviour of young children, reactions of autistic children to different social and material environments, and the distinctive behavioural features of 'brain-damaged' children. In this work they present some of the fruits of their experience of that type of research in which the spontaneous behaviour of human subjects, or selected categories of their actions and reactions, are systematically recorded by trained observers, these records constituting the data for scientific analysis. For more than a century students of human development have pursued such methods by diary records, by recorded commentary, by various forms of systematic sampling, and the like; this book is largely devoted to giving a most useful and expert account of modern methods, techniques and equipment. Chapters deal with tape-recording, check-lists and event-recorders, and film and videotape. These and other sections contain good examples of animal and human studies; most of the latter deal with children and some of these are detailed accounts of the authors' research projects at the Human Development Research Unit, Park Hospital for Children, Oxford. No research psychologist intending to employ these techniques can afford not to consult this work for its information, illustrations and references. A chapter on some statistical, logical and mathematical models for the sequential analysis of elements of behaviour demands fairly expert acquaintance with the topics. There is an appendix, Glossary of Motor Patterns of Four-Year-Old Nursery School Children, by W. C. McGrew, Department of Psychology, Edinburgh. The work of some of those currently engaged on similar research in this country, for example Chance and his colleagues at Birmingham University, receives attention, but although Blurton-Jones appears in the author index, I did not notice reference to him in the text.

This is an idiosyncratic book and it is all the better for that; in every way it is unlike most of the texts, handbooks and symposia in psychology and allied fields that are beginning to constipate the library shelves. This is true even of its physical appearance: it is handsomely produced with good illustrations, but for that one must pay rather a lot. But like every real personal encounter it induces mixed feelings, including exasperation. The work appears as a monograph in The Bannerstone Division of American Lectures in Living Chemistry. It is not reported that it was originally a series of lectures but one suspects this because some passages are like those provocative statements of half-truths, or even smaller fractions, that serve to shock a particular audience out of accepted views. In permanent print, however, they may only mislead the ill-informed and antagonize the others.

It may be good for us to be castigated by Tinbergen for our yearning for scientific respectability through a premature fashion for controlled experiment. But it is not historically correct to explain the decline in observational studies by that motive alone. Even the vicissitudes of research economics played a part in that decline. The authors make good use of the survey articles by Arrington and by Gellert; less than he deserves of Wright's chapter in Mussen's *Handbook*. But they give insufficient attention to the crisis that followed the high output of descriptive studies in the 1930s, to which Gellert referred in her conclusion that 'the most significant development...is the trend...towards a selection of specific, theory-determined aspects of behavior for study, with a view to applying results to the testing of stated hypotheses, or to answering a particular question'. Indeed the authors' own uses of the methods of investigation they so ably espouse illustrate that 'significant development'.

'Psychology', in contrast to ethology, we are told, 'has largely restricted itself to well-controlled laboratory experiment'. Even in a polemic, that is too exaggerated. 'While fundamentally more complex, the behavior of *Homo sapiens* can in principle be treated in the same way as that of any other animal studied in his natural habitat' is a declaration of faith; but in practice it is not a guide to action for all types of psychological inquiry. What is man's 'natural habitat'? In the editor's foreword and in the publisher's announcement the term 'ethological' is given particular emphasis. In their concluding chapter the authors cite, with an approval that most psychologists will share, Tinbergen's description of the ethological approach in biology. But few of the examples of systematic observation of human behaviour in this book are ethological in

the traditional sense; we lose a valuable term too soon in the erosion of language if all non-experimental field observations are to be attributed to ethology.

There are some points of dating and attribution which are distinctly unconventional. In their account of Bridges' study of emotionality in infants they comment on her classification of thumb-sucking as a pleasure-seeking activity: 'may it not frequently occur as a displacement activity?' In his famous monograph, published in the same year as Bridges' work, von Uexkull uses the term *Übersprungbewegung*, which has been translated as 'displacement activity', but was not yet the concept which Tinbergen named. And surely the association of the term 'hyperkinetic' with a syndrome of disturbed behaviour in childhood antedates 1955 by many years: see Bender (1949), for example.

So much for the exasperations. Apart from the exposition of methodology there are many fascinating things reported from the authors' own work. For example, the demonstration that the differences in the duration of visual fixation in 'hyperkinetic' brain-injured children and others are situationally determined; and for another, the contrast in the strategies between the same groups in their sampling of the environment.

C. J. PHILLIPS

Theories of Development. By JONAS LANGER. New York: Holt, Rinehart & Winston. 1969. Pp. xi+191. $5.95.

Professor Langer has written a somewhat uneven book. It is uneven both in the relative space given to the categories under which he subsumes the theories he examines and in the level of analysis which he offers.

The brief introductory chapter discusses concepts of change, models of man and research methods concisely and clearly. It also introduces Langer's plan, which is to categorize theories under three headings: Psychoanalytic Theories, Mechanical Mirror Theories ('focusing on continuous quantitative accumulation of response elements that are stimulated by the environment'), and Organic Lamp Theories ('the primary focus of which is the organism's construction of its own rational systems of action').

The chapter on psychoanalytic theories is done in the same objective and balanced manner as the introduction; though it is too slight to be more than an outline summary, it does take into account some of the more recent developments. The next chapter is on mechanical mirror theories and here the tone changes, appearing to become both less objective and more critical. It is not always easy to see the real point of some of Langer's strictures: they appear on occasion to be open to the very objection that they are often making, namely that the conclusions are not warranted by the evidence.

The chapter on organic lamp theories is twice the length of either of the previous chapters, though it remains expository rather than evaluative. It draws heavily on Werner, to whom the book is dedicated, and at points, such as in the discussion of genetic epistemology, the chapter becomes a straight outline of Werner's views. Langer gives a useful account of Piaget's work on cognitive development and here his ability to condense a sustained argument shows at its best. This discussion of organismic theories is much more detailed than either of the previous summaries but, as with the rest of the book, it is difficult to envisage the particular audience to whom it is addressed.

In his preface Langer suggests that the reader should read his final chapter before tackling the text so that he might at least know 'my general perspective before "accepting" my formulation and evaluation of developmental problems and theories'. This could be advice of mixed value. Langer's 'general perspective' is not difficult to pick up from the text and, as the last chapter is in some ways the least satisfactory in the book, it may well put the reader off what are, in spite of the foregoing comments, useful summaries of some major positions. In this final chapter Langer offers his own notes towards a comprehensive theory, and these I found unsatisfactory. He specifies two major tasks facing the theorist. The first is 'to specify the formal determinants of the end (most advanced and mature) stage that will be achieved in life and the direction that development must take in order to reach that final stage. When this is done it should become possible to postulate tentatively the functional structures that must make up the organization of the initial (most primitive) stage if development is to take the course specified. The second theoretical task is 'to specify the generative rules that must be built into the organi-

zation of functional structures of the initial stage if this stage is to determine the direction of development'. Langer then suggests that logical implication is 'at present the most appropriate mode of explaining how new stages for functional structures and systems of action are generated'. Though logical implication may be a necessary quality of any explanation I am not clear how it can explain adequately the generation of anything (and if it could, the 'at present' would not be necessary).

Langer sees the tasks he proposes as having 'some resemblance' to what Aristotle called final determination: but final determination for Aristotle operated only in conjunction with the other three modes of 'causality', material, formal and efficient. It is, however, unfair to press this point, as Langer himself admits that the limits of the book do not permit an exhaustive discussion of his views. Perhaps this is the real weakness of the book: it is too brief to be much more than an outline of selected theories but it repeatedly raises points that are not sufficiently developed to be more than distracting. It could prove a useful revision text for third-year undergraduates.

W. M. TIDMARSH

Minnesota Symposia on Child Psychology. Volume 3. Edited by JOHN P. HILL. London: Oxford University Press. 1969. Pp. viii + 163. 48*s*.

The intending research student in child development cannot afford to miss this book. It provides models of research strategy, workmanlike execution of projects, conceptualization, and, possibly most important of all, communication. This lucidity and the compelling nature of the material ensure a wide readership at all levels amongst those interested in the behavioural sciences. It is in many ways a difficult book to read; difficult because it stimulates so many ideas one would like to pursue that to persist in further reading seems pointless when there is so much to be done.

This is the third volume in a series based on papers from the Annual Minnesota Symposia on Child Psychology. It more than lives up to the promise of the previous two. The material here is based on six papers presented at the 1968 Symposium when researchers of some eminence were invited to give overviews of their work. This volume does much to add to their stature.

It is an invigorating collection of papers and leaves a clear impression that here at last are a group of workers who are approaching the real problems in child development and are not afraid to go back to first principles. Gordon Cantor ends his piece on stimulus familiarization with admitted ingenuousness when he says: 'Older children respond to the world differently from younger children, at least partly, because they have simply been looking at it longer'. Victor Denenberg in his paper provides further justification for the current interest of developmental psychologists in infancy. Maccoby introduces her topic of the development of stimulus selection thus: 'The problem of stimulus selection arises from the fact that there is a greater array of stimulation potentially available to an organism at any given moment than the organism can take in, remember or respond to.' The Sherifs, presenting a paper on adolescent groups, say, 'People *do* act differently when they know they are being watched, this is particularly true of adolescents who are on their guard in relations with adults.' Here are several researchers who state the apparently obvious, but not only do they produce their own research findings but also a fairly comprehensive survey of related investigations and theories which demonstrate to us that the obvious has more to it than mere 'face validity'.

Three of the papers deal with attentional processes: Maccoby and Cantor on stimulus selection and familiarization respectively and Jeffrey on early stimulation and cognitive development. Cantor finds that non-familiarized stimuli are responded to more quickly, projected for longer and preferred to the familiarized. His most interesting proposition centres around the problem of boredom. Jeffrey, like Cantor, returns to habituation in conceptualizing the effects of varying regimens of exposure to stimuli. Maccoby, in a masterly paper, uses data from her research on dichotic listening as a reference point to develop many hypotheses about the development of stimulus selection and skilfully relates these to contemporary thinking in the field.

The remaining papers deal with several other important areas. Denenberg's analysis of the implications for human development of animal studies of early experience must be commended for its clear thinking and neat exposition. He demonstrates that early experience can modify characteristics thought to be least amenable to environmental modification and also the

lasting effects of such experience. Adolescence, although increasingly a fashionable topic, is still thinly researched. Muzafer and Carolyn Sherif, from the report of their work, have done much to cure the fault. Their most exciting contribution, however, is methodological in that they have conducted an observational study in the field. Sutton-Smith and Rosenberg's paper on 'Modeling and Reactive Components of Sibling Interaction' also has methodological significance. It is a tidying-up operation which attempts to eliminate the disadvantages of previous studies, such as inadequate controls, ignoring the sex of siblings, spacings, ordinal positions, etc. It is an attempt which, although necessarily complex, proves immensely worthwhile.

The University of Minnesota is to be congratulated on this significant contribution to the literature in child psychology.

MARY CROXEN

Behaviour and Adaptation in Late Life. Edited by EWALD W. BUSSE and ERIC PFEIFFER. Boston: Little, Brown. 1970. Pp. xii + 395. $13.50.

Duke University, North Carolina, has been one of the pioneers in work on ageing, with its Center for the Study of Aging and Human Development. In this book, 17 authors, 15 of them from Duke, have written about some aspects of gerontology in their own disciplines. Thanks to the careful ordering and interrelation of their chapters, a picture emerges of the aged person as a resultant of many forces ranging from physical health to social convention. 'This book is written for those with an interest in any of a number of aspects of aging', say the editors, but there is clearly a medical bias, for terms such as 'ischemia', 'hypoxia' and 'necrosis' occur without explanation in the otherwise simple and readable text.

Age changes which result from innate causes or from physical or physiological trauma can be discussed independently of cultural background, as in the chapter on 'Theories of Aging', but methods of dealing with those changes vary from country to country. However, American ways can serve as a model (or an awful warning) and are therefore worth noting. From the many chapters on medical care, hospitalization and social services, it is clear that America is not a place for the aged at present. There is a shortage of suitable facilities, with lack of provision for improvement, and only a grudging acceptance of the need for financial aid to the aged sick. However, it would be unwise to assume that our own society is very much more successful in providing care and services and finance for the aged, since a pessimist could have written many similar passages about the aged in Britain.

The theoretical aspects of finance are considered in two chapters on economic theories and practices. In 'Economics of Retirement' a contrast is drawn between the effects of cutting short the working life by lowering the retirement age or by shortening the working week. If people retire younger, a larger retired group are left without economic bargaining power and without means of increasing their income in times of inflation. However, a shorter working week with late retirement may have unfortunate consequences for those who 'age' young. In 'The Aged and Public Policy' it is pointed out that those over 65 could bargain through political pressure, since there are enough in this age group to affect party fortunes. (There seems to be a lesson here for the elderly in Great Britain too!) But many very old people are too frail to organize themselves into a pressure group, and younger people are often not willing to contemplate their own possible old age and sickness.

Two chapters deal with psychological changes with age. 'Sociological Aspects of Aging' reviews not only how the aged fit into American society but also briefly covers such topics as alteration in mental abilities, personality changes and the controversy as to whether disengagement from personal involvements and activities occurs in a voluntary fashion or as a result of circumstances which prevent opportunities for activity. Evidence for the latter view is given. 'Intellectual and Cognitive Changes in the Aged' is narrow in focus and mentions some recent American research only. However, it does make clear that the assessment of mental ability is complicated, and that the factors underlying deterioration are complex.

There are also two chapters concerned with the psychiatry of old age. 'Functional psychiatric Disorders in Old Age' provides a grim picture of how little is done to relieve such disorders in the U.S.A. at the moment, though it presents an optimistic view of what could be done. In contrast, 'Organic Brain Syndromes' makes interesting but depressing reading because there is not only difficulty in exact diagnosis, but also no present hope of reversing deteriorations based on loss of

brain cells. The changes in brain tissue which underlie organic syndromes are well described in a further chapter, 'The Brain and Time'.

Psychologists already interested in ageing are likely to have encountered fuller expositions elsewhere of their immediate fields of interest. One suspects that physiologists, neurologists and psychiatrists will say the same of their own areas. However, the virtue of this book lies in the fact that many disciplines are represented and that their interdependence is made clear. The book, backed by other more specialized material, might possibly help to persuade a few more people to specialize in the gerontology of their own discipline. SHEILA CHOWN

Television and Delinquency. By J. D. HALLORAN, R. L. BROWN and D. C. CHANEY. Leicester: Leicester University Press. 1970. Pp. 221. 30*s*.

This book, which sets out to show the relationship between television viewing and delinquency, is the third working paper to be published by the Television Research Committee. It differs from previous working papers in that the work was actually initiated by the Committee itself and was completed at the Centre for Mass Communication Research, which the Television Research Committee established at Leicester in 1966.

Criminologists and social psychologists will already be familiar with investigations into the relationship between elements of the media and delinquency; they may well be prejudiced against them. For the unfortunate truth is that such studies have often not sprung from any intrinsic intellectual problem encountered by academics, but have rather been foisted upon social scientists by research bodies, typically following in the wake of public outcries about certain aspects of the media. As Kingsley Davis pointed out some time ago, there is no more popular study than that which sets out to provide confirmation for an 'evil causes evil' hypothesis. Whenever radio, gangster films, horror comics, violent television and open-air pop concerts come under attack, one can expect to see the attack strengthened by the suggestion that they are critically associated with other regrettable phenomena—illicit sex, drug-taking, juvenile delinquency.

Halloran and his colleagues break away from this unfortunate tradition in an important way. They adequately demonstrate in the long theoretical prologue to their research the very serious methodological and theoretical limitations of most media delinquency studies which have preceded their work. In this section they show themselves to be not merely sophisticated about the way in which the media operate—they make full use of the contemporary stress upon the different functions which the television may serve for different groups—but they also give highly sensitive interpretations of the delinquent side of the issue. They draw upon the work of the important American school to show how theories about the impact of the media can be brought into alignment with Mertonian anomie formulations, the subcultural theory of Cohen, Cloward and Ohlin, and the lower-class cultural theory associated with Walter Miller.

This is all very welcome. Unfortunately, this sophistication is not quite carried over into the empirical work. The authors are quite prepared to admit this imbalance between theory and practice. They are consistently self-deprecating. They might, however, have produced a more satisfying study had they simply concentrated upon a few elements of the viewing situation rather than going in for the type of conventional analysis which was favoured by their less-sophisticated predecessors in the field.

Basically, the study consists of a comparison between three groups of young people: probationers, a matched sample of non-probationers and a lower middle-class group. The second half of the book consists of the tabular presentation of the significant differences in viewing habits between such groups. The researchers found that the delinquent group, the probationers, expressed a disproportionate interest in 'exciting programmes', that this group talked considerably less about their viewing than the other two samples, that they were particularly aware of hero figures in television programmes and that they were more likely to mention pop performers than members of the other two samples. The authors do not for one moment suggest some causal link between these particular viewing attributes and indulgence in delinquency; the furthest they wish to go is to suggest that television may play a minor contributory role in the production of delinquent behaviour.

The least satisfactory part of the book comes in the last few pages, where the authors attempt to relate in a meaningful way the delinquent's viewing habits to his actual delinquency. They

suggest, for example, that 'those who watch television but never or seldom talk about it may accept its suggestions in a far less critical fashion'. This sounds like an interesting hypothesis to investigate, but it is a dubious way of giving significance to their findings. It can surely be just as readily argued, following social psychological evidence, that group discussion of attitudes and ideas can play an important part in not just confirming these as part of the individual's attitudinal repertoire, but also in leading him to realize them in behaviour. There is no room to go into the other objections to this final section. But then one has little doubt that Halloran and his associates are already familiar with them.

The penalty which the authors pay for their sophisticated theoretical indulgence in the first part of the book is that they disappoint the reader when he encounters their research in the second section. Had their concerns been less global—and one notes with approval in this respect the recently published study by Halloran, Elliott and Murdoch on the 1968 anti-Vietnam War demonstration—then they might have had more opportunity to get to grips with the subtleties of media influence instead of reworking, albeit in a very thoroughgoing manner, a type of study whose methodological and theoretical limitations were already evident in the somewhat meagre advances it had occasioned in the past.

LAURIE TAYLOR

Behaviour Therapy in Clinical Psychiatry. By V. MEYER and E. S. CHESSER. Harmondsworth: Penguin Books. 1970. Pp. 288. 15*s*.

In the last few years there has been an increasing number of books and selected readings describing behaviour modification and behaviour therapy. We are approaching the stage when we look more critically at each newcomer to see if it justifies its place. This book easily passes most tests; it is concise, readable and discusses a wide range of techniques in adequate detail. There is an extensive reference list.

No book is free of faults. The introductory theoretical section is rather weak and it is irritating to find the description of discrete trial discriminative avoidance training offered as a general definition of avoidance learning. In places the authors adopt a rather telegraphic style; this undoubtedly reflects an attempt to remain within their allotted space, but it has the effect of polarizing controversies to an extent that makes them look like caricatures. This effect is perhaps most marked when the authors discuss operant conditioning techniques and cognitive factors in behaviour therapy. The authors have a preference for cognitive explanations of human behaviour; this preference is perfectly legitimate and is well represented in the literature of behaviour therapy, but at times they appear to equate attention to verbal instructions and verbal report with the necessity for postulating crucial mediating cognitive process. Their discussion of this area would benefit from more rigour, but I suspect such an extension of their discussion may have been impossible because of limited space.

These comments do not detract from the firm opinion that this book represents a very concise and useful contribution among introductory textbooks on behaviour therapy. It can be recommended to trainee clinical psychologists, trainee psychiatrists and others wishing to learn more than the rules-of-thumb of behaviour therapy.

J. T. QUINN

Rehabilitation Research. By GEORGE N. WRIGHT and ANN BECK TROTTER. Madison: University of Wisconsin. 1968. Pp. xiii + 674. $14.00.

This lengthy volume is a review of research in vocational rehabilitation the purpose of which is 'to disseminate information on an important portion of the Research and Demonstration Program administered by the U.S.A. Federal Vocational Rehabilitation Administration'. The authors report 97 separate studies which have been selected to be useful to rehabilitation agencies in planning and staff development, to universities in research and student seminars, and to other persons interested in the field of rehabilitation. Whether they will or will not be useful to the second of these categories is rather dubious. Nevertheless, much of the material (in a volume which can be extremely tedious to read through and which should only be dipped into at relevant chapters) is evidence of careful and painstaking work of value to psychologically orientated rehabilitation staff.

The key problem with this book is that the separate studies, reasonably enough reported under 20 chapter headings such as Cardiovascular Diseases, Cerebral Palsy, Mental and Personality Disorders, Speech and Hearing Disorders, Homebound Disabled, Ageing and Chronic Illness and so forth, and each prefaced by a brief description of the problem area written apparently at the level of first year undergraduate or third year student nurse, are in effect reports of reports. The authors have surveyed the work of the researchers involved in the V.R.A. programme throughout the U.S.A., and while acknowledging the original authorship of the papers have 'tidied up' the presentation by a uniform layout of each study under the headings of Methodology, Results, Discussion and Summary. The writing is therefore very compressed and although all the basic data underlying conclusions are presented there is often a feeling that the bulk of related studies outside the V.R.A. programme has been ignored and that the work has gone on in a kind of academic enclave. This may not in fact be so but the lack of any references other than to the sponsored projects tends to endorse this view.

What one cannot but be impressed with, however, is the vast amount of specialist time and resources poured into these researches. Clinical and experimental psychologists will, one feels, labour under a feeling of irritation at having research reports edited for them with unexplained assumptions and skimping of details of experimental design, but occupational therapists, rehabilitation supervisors and psychologists working in this applied field will all wish at least to have the volume available and to select, within their own areas of interest, from the wealth of data and rehabilitation experience here made available.

D. F. CLARK

Your Child is Asleep: Early Infantile Autism—Etiology, Treatment, Parental Influences
By Austin M. DesLauriers and Carole F. Carlson. Homewood, Ill.: **Dorsey Press. 1969. Pp. x + 403. 67*s*. 6*d*.**

This immensely readable book proposes an interesting new model to explain early infantile autism. Autistic behaviour is assumed to be the result of a form of sensory and affective deprivation which derives from a functional imbalance between the arousal systems associated with the reticular formation and the limbic system. The authors draw heavily on Routtenberg's work, which has provided evidence for assuming that motivational state depends on a dynamic equilibrium between a drive-response system (the reticular formation) and incentive-motivation system (limbic system). Autistic children are assumed to have such a high threshold of affective arousal (due to some malfunctioning of the limbic system) that the effects of the limbic arousal system in moderating and modulating behaviour triggered off by the drive mechanisms involving the reticular formation are inhibited, so that the possibility of effectively reinforcing any behavioural response is absent. (The model thus provides an explanation of the failure of autistic children to learn.) Therapy itself is aimed at correcting the imbalance between these two systems by activating the limbic arousal system through the use of stimulation which elicits marked affective behaviour. In the context of the heightened emotional involvement of the child with the therapist (and later with parents and siblings) learning can then take place.

DesLauriers and Carlson present their theory very persuasively, but it is singularly unfortunate that their work is based on only five case histories. The success they report with these children is encouraging, but, as they themselves point out, such changes do not warrant any claim to the validity of the neurophysiological model proposed. However, since their model does provide a coherent rationale for research and therapy the book is essential reading for all research workers interested in autism and the general field of developmental arrest. The model may also prove of value in providing a theoretical framework for viewing autism, not as a disease entity, but a subspecies of a broad spectrum of behaviours which are characterized by persistent failure to learn.

ANDREW TREACHER

Pavlovian Approach to Psychopathology: History and Perspectives. Edited by W. H. GANTT, L. PICKENHAIN and CH. ZWINGMANN. Oxford: Pergamon Press. 1970. Pp. 341. 84s.

The growth of behaviour therapy has brought forth a large number of books devoted to the application of Pavlovian and operant conditioning to the treatment of neurotic and psychotic patients. Most of these tend to be written or edited by a small group of behaviour therapists and to contain pretty much the same material with slight variations in arrangement and theoretical content. The book of readings edited by Gantt and his colleagues is not entirely different from the usual run of these, containing for instance a reprint of Wolpe's 1952 paper and a reprint of Lindsley's paper on 'Characteristics of the behaviour of chronic psychotics as revealed by free-operant conditioning methods'. But the majority of papers are rather different, emphasizing a more strictly Pavlovian approach than is usual, and stressing contributions by Russian authors and by Gantt and his school.

There are three parts to the book. Part I, entitled 'History and Foundations', contains six papers; Part II, entitled 'Physiological Systems', contains five papers, and Part III, entitled 'Experimental Neuroses and Clinical Applications of the Conditioned Reflex Method', contains seven papers. The choice of papers makes one wonder for whom the book is intended. Some of the contributions, such as Bykov's 'Conditional reflex connections of the kidney', are clearly intended for a knowledgeable and sophisticated audience; such an audience, however, would surely be familiar with Broadhurst's classical paper on 'Abnormal animal behaviour', reprinted from *The Handbook of Abnormal Psychology*. On the whole I can perhaps say that the choice of papers included, and the point of view taken, are sufficiently unlike those found in other books to make this a worthwhile contribution, useful perhaps more to the academic lecturer and research worker than to the practising psychiatrist and clinical psychologist. There has been a relative neglect of Gantt's empirical and theoretical contributions, particularly his notions of schizokinesis and autokinesis; this book may do something to redress the balance. Unfortunately, it does not contain what one would have very much liked to see, namely a detailed discussion of the differences between the Pavlovian approach and that of other learning theorists. These differences are numerous and far-reaching but they are seldom discussed in such a manner as to clarify the issues involved. This book, in singling out the Pavlovian approach, should surely have pointed the contrast rather than simply include such papers as that of Wolpe, who is much more a Hullian than a Pavlovian. Perhaps the authors hoped that readers would be stimulated to formulate their own thoughts on this topic; even so, one would have preferred more detailed discussion by someone like Gantt, who after all had the privilege of working with Pavlov.

H. J. EYSENCK

Research Concepts in Human Behavior: Education, Psychology, Sociology. By G. C. HELMSTADTER. New York: Appleton Century-Crofts. 1970. $7.50.

The presentation, for pedagogic purposes, of research methodology (or what is involved in research) is an appallingly difficult task—in general far more difficult than the performance of any individual experiment. G. C. Helmstadter approaches the problem by expounding concepts common to the three areas of experimental psychology, educational research and empirical sociology, and by noting the 'task demands' involved in research, such as the storage and retrieval of information. He accomplishes this by maintaining an empiricist bias, as we shall see. The basic fault is indicated in the title of the book—research *concepts* are divorced from their theoretical articulation.

The book begins with 40 pages on 'The Nature and Function of Research': a brief sketch of 'the activity of solving problems which leads to new knowledge using methods of inquiry which are currently accepted as adequate by scholars in the field'. Next come 90 pages on technique—descriptive, Skinnerian, questionnaire etc., followed by 30 pages on the storage and retrieval of information, and 250 pages on quantification. Finally, there is a 50-page epilogue on the techniques of communicating and implementing results.

As in most texts designed for course work there are two strands to the presentation. One is

a simplified account of the principles and theory of full-blown research—a sort of preview, as it were, of what scientists are like. The other is an introduction to the strategy and tactics of research. This is a means of easing students into a situation where they will be faced with problems of methodology, armed, hopefully, with a few good tips which will aid them in designing investigations and in making extended judgements on related investigations to which they are less committed than to their own. Helmstadter's work leans heavily towards this second aspect of instruction. The major indication of this bias consists in the omission of any reference to men who, by example or recommendation, have raised fundamental problems of scientific methodology in the fields considered: Craik, Durkheim, Feyerabend, Freud, Kuhn, Marx, Merton, Piaget, Weber, etc. Thus the theoretical principles underlying Helmstadter's work are not made explicit for the student's consideration. The nature of these implicit principles may be gathered from the following examples.

On p. 65 we are told that there is a natural (*sic*) trend towards more experimental approaches as knowledge increases in a field, and the arguments adduced for this are technical (p. 118): the experimental method permits control, manipulation and simplification of situational variables. Certainly, but what are our criteria for judging the value of these factors and their supposed incarnation in a given experiment? Our suspicion that Helmstadter has reversed the relationship between theory and application is confirmed on p. 221, where the non-congruence between correlation and causal relationship is presented as an *additional* point to remember about correlations. It is true that not every statistical phenomenon represents a scientific fact, but the lack of an articulated theory cannot be patched up by a simple technical memorandum.

In fact, as is by now obvious, the author's *forte* is the handling of technical matters (although, rather oddly, there is no analysis of non-parametric methods), in particular of certain aspects of quantification. Even so, the sorts of issues raised by such works on quantification as the 1961 Isis symposium are summarily dealt with in a few pages after the section on 'Statistical Reasoning'. Again technical factors are given priority over theory to an unwarranted degree. One particularly obscure feature of the discussion is the following argument. Numbers are basically value-free, evaluation succeeds quantification, and 'in no sense should measurement be considered evaluative in nature' (p. 263). To draw so simple a distinction between quantification and evaluation avoids the mistake of conflating them, but seems also to rule out any possibility of understanding their interrelationships.

To conclude, there is no question but that the useful hints contained in the book are vitiated by its low-level theory. There is no sustained analysis of the fundamental differences between the methodologies of psychology, sociology and educational research. The elements common to these three fields turn out to be the lowest common denominators involved in the conceptualization of experimentation. These are serious limitations. Students reading this book are offered little more than a technician's substitute for a foundation of knowledge.

N. H. FREEMAN

The Nervous System. By PETER NATHAN. Harmondsworth: Penguin Books. 1969. Pp. 391. 10*s*.

Dr Nathan has produced a comprehensive and attractively written book. It is at its best when dealing with physiological and neurological topics and at its weakest when dealing with more psychologically orientated questions. The book is essentially aimed at the general reader but would be of value to first-year undergraduates requiring introductory reading material. More advanced students will nonetheless find the book stimulating, but the absence of specific references and a bibliography is annoying. It is also surprising that although the author is a neurologist and draws heavily on comparative material there is no chapter specifically on the evolution of the nervous system.

ANDREW TREACHER

Br. J. Psychol. (1971), 62, 1, *pp.* 143–145
Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review)

ADELSON, D. & KALIS, B. L. (eds.). *Community Psychology and Mental Health: Perspectives and Challenges.* London: International Textbook Co. 1970. Pp. vii + 341. £3.80.

ALLEN, G. W. *William James.* Minneapolis: University of Minnesota Press. 1970. Pp. 48. 42½ np.

ANDERSON, J. R. L. *The Ulysses Factor: the Exploring Instinct in Man.* London: Hodder & Stoughton. 1970. Pp. 352. £2.50.

AUSTRALIAN BROADCASTING CONTROL BOARD. *Attitudes to Television.* Melbourne: Australian B.C.B. 1970. Pp. 112. Price not given.

BADIA, P., HABER, A. & RUNYON, R. P. *Research Problems in Psychology.* London: Addison-Wesley. 1970. Pp. 349. £2.60.

BLACK, J. *The Dominion of Man: the Search for Ecological Responsibility.* Edinburgh: University of Edinburgh Press. 1970. Pp. vi + 169. £1.50.

BLATT, B. *Exodus from Pandemonium: Human Abuse and a Reformation of Public Policy.* Boston, Mass.: Allyn & Bacon. 1970. Pp. xix + 268. Price not given.

BRUMLIK, J. & YAP, C.-B. *Normal Tremor: a Comparative Study.* Springfield, Ill.: Thomas. 1970. Pp. xiii + 93. $7.50.

CRAFT, M. *Speech Delay: its Treatment by Speech Play.* Bristol: Wright. 1969. Pp. 112. 70 np.

CREWS, F. (ed.). *Psychoanalysis and Literary Process.* Cambridge, Mass.: Winthrop. 1970. Pp. viii + 296. £2.00.

DENNY, M. R. & RATNER, S. C. *Comparative Psychology: Research in Animal Behavior.* Homewood, Ill.: Dorsey Press. 1970. (Revised edition.) Pp. xiv + 869. $12.50.

DODWELL, P. C. (ed.). *Perceptual Learning and Adaptation.* Harmondsworth, Middx.: Penguin Books. 1970. Pp. 502. 60 np.

DONCHIN, E. & LINDSLEY, D. B. (eds.). *Average Evoked Potentials: Methods, Results, and Evaluations.* Washington, D.C.: NASA. 1969. Pp. xviii + 400. $2.00 (paper).

DOUGLAS, M. *Natural Symbols: Explorations in Cosmology.* London: Barrie & Rockliff. 1970. Pp. xvii + 177. £2.25.

DOUGLAS, M. (ed.). *Witchcraft: Confessions and Accusations.* London: Tavistock Publications. 1970. Pp. xxxviii + 387. £3.15.

EASSON, W. M. *The Dying Child: the Management of the Child/Adolescent who is Dying.* Springfield, Ill.: Thomas. 1970. Pp. vii + 103. $5.75.

EDELSON, M. *Sociotherapy and Psychotherapy.* London: University of Chicago Press. 1970. Pp. xviii + 266. £5.40.

ELKIND, D. *Children and Adolescents: Interpretive Essays on Jean Piaget.* London: Oxford University Press. 1970. Pp. xi + 160. £1.00.

EVANS, R. I. & ROZELLE, R. M. *Social Psychology in Life.* Boston, Mass.: Allyn & Bacon. 1970. Pp. xv + 359. Price not given.

FLANIGAN, P. J., BAKER, G. R. & LAFOLLETTE, L. G. *An Orientation to Mental Retardation: a Programmed Text.* Springfield, Ill.: Thomas. 1970. Pp. xiv + 203. $9.50.

FRAISSE, P. & PIAGET, J. (eds.). *Experimental Psychology: its Scope and Method. Vol. IV.* (Le Ny, J.-F., de Montpelier, G., Oleron, G. & Flores, C.) *Learning and Memory.* London: Routledge & Kegan Paul. 1970. (First published in French, 1964.) Pp. viii + 376. £4.00.

FREDERICK, C. J. *The Future of Psychotherapy.* Boston, Mass.: Little, Brown. 1969. Pp. xii + 415. £3.75.

Geiwitz, P. J. *Non-Freudian Personality Theories.* Hemel Hempstead: Prentice-Hall. 1969. Pp. xi + 147. £1.00.

Gergen, K. J. *The Psychology of Behavior Exchange.* London: Addison-Wesley. 1970. Pp. xi + 109. 85 np.

Hamburg, D. A. (ed.). *Psychiatry as a Behavioral Science.* Englewood Cliffs, N.J.: Prentice-Hall. (Spectrum paperback.) 1970. Pp. xii + 114. $1.95.

Herrington, R. N. (ed.). *Current Problems in Neuropsychiatry: Schizophrenia, Epilepsy, the Temporal Lobe.* Ashford, Kent: Headley Bros. 1970. Pp. 184. £1.60.

Hilgard, J. R. *Personality and Hypnosis: a Study of Imaginative Involvement.* London: University of Chicago Press. 1970. Pp. x + 304. £5.65.

Koskiaho, B. *Level of Living and Industrialisation.* Jyväskylä, Finland: Jyväskylän Yliopisto. 1969. Pp. 102. Price not given.

Kraus, E. A. *Pathways Back to the Community.* New York: Springer. 1970. Pp. 40. $3.50.

Laing, R. D. *Knots.* London: Tavistock Publications. 1970. Pp. 90. £1.50.

Legge, D. (ed.). *Skills.* Harmondsworth: Penguin Books. 1970. Pp. 378. 50 np.

Levinson, B. M. *Pet-oriented Child Psychotherapy.* Springfield, Ill.: Thomas. 1969. Pp. xxiii + 202. $9.75.

Lingeman, R. R. *Drugs from A–Z.* London: Allen Lane The Penguin Press. 1970. (Originally published in U.S.A., 1969, McGraw-Hill.) Pp. 262. £2.50.

Love, J. O. *Worlds in Consciousness: Mythopoetic Thought in the Novels of Virginia Woolf.* London: University of California Press. 1970. Pp. xvi + 268. £3.80.

Lucas, R. A. *Men in Crisis: a Study of a Mine Disaster.* London: Basic Books. 1969. Pp. ix + 335. £4.65.

Luthe, W. *Autogenic Therapy, Vol. IV: Research and Theory.* New York: Grune & Stratton. 1970. Pp. xi + 276. $14.75.

McNeil, E. B. *The Psychoses.* Hemel Hempstead: Prentice-Hall. 1970. Pp. x + 194. £1.40.

Maas, P. *The Valachi Papers.* London: Panther Books. 1970. (First published, 1968.) Pp. 209. 30 np.

Maclay, D. T. *Treatment for Children: the Work of a Child Guidance Clinic.* London: Allen & Unwin. 1970. Pp. 247. £3.00.

Marrow, A. J. *The Practical Theorist: the Life and Work of Kurt Lewin.* New York: Basic Books. 1970. Pp. xxii + 290. £4.15.

Mental Health Book Review Index. *An Annual Bibliography of Books and Book Reviews in the Behavioral Sciences.* Vol. 15. New York: New York University. 1970. Pp. 88. $10.

Paranjpe, A. C. *Caste, Prejudice and the Individual.* Bombay: Lalvani. 1970. Pp. 236. Rs. 24.00.

Piaget, J. *The Child's Conception of Time.* London: Routledge & Kegan Paul. 1969. (First published in French, 1927.) Pp. xi + 285. Price not given.

Rapoport, R. N. *Mid-career Development: Research Perspectives on a Developmental Community for Senior Administrators.* London: Tavistock Publications. 1970. Pp. xiii + 290. £2.90.

Ratcliffe, T. A. *The Child and Reality.* Hemel Hempstead: Allen & Unwin. 1970. Pp. 141. £2.00.

Rauhala, L. *Die Seinsfrage in der sog. neueren Psychosomatik.* Helsinki: Turun Yliopisto. 1970. Pp. 32. Price not given.

Rhine, L. E. *Mind over Matter.* London: Collier-Macmillan. 1970. Pp. x + 402. $7.95.

Ricoeur, P. *Freud and Philosophy: an Essay on Interpretation.* London: Yale University Press. 1970. Pp. xiv + 573. $15.00.

Robinson, D. N. (ed.). *Heredity and Achievement: a Book of Readings.* London: Oxford University Press. 1970. Pp. x + 444. £1.25 (flexi).

Roff, M. & Ricks, D. F. (eds.). *Life History Research in Psychopathology.* Minneapolis: University of Minnesota Press. 1970. Pp. vii + 321. £4.75.

SCHMIDEBERG, M., RUSSELL, D. H. & WHISKIN, F. E. *Massachusetts Court Clinics.* London: International Journal of Offender Therapy. 1970. Pp. 20. $1.75.

SOUTHWELL, E. A. & FELDMAN, H. (eds.). *Abnormal Psychology: Readings in Theory and Research.* Belmont, Calif.: Brooks-Cole. 1970. Pp. ix + 389. Price not given.

STERN, G. G. *People in Context: Measuring Person-Environment Congruence in Education and Industry.* Chichester: Wiley. 1970. Pp. xxvi + 402. £6.15.

TREDGOLD, R. & SODDY, K. *Mental Retardation* (Eleventh edition.) London: Baillière, Tindall & Cassell. 1970. Pp. xi + 490. £5.00.

VERNON, M. D. (ed.). *Experiments in Visual Perception.* (Second edition.) Harmondsworth: Penguin Books. 1970. Pp. 472. 50 np.

VILJANEN, E. *The Level of Basic Education in relation to the Formation of the Development Milieus of Primary School Teachers.* Jyväskylä, Finland: Jyväskylän Yliopisto. 1970. Pp. 280. Price not given.

WEISKRANTZ, L. *Inroads and Detours in Psychology.* Oxford: Clarendon Press. 1970. Pp. 28. 30 np.

WEST, D. J. (ed.). *Criminological Implications of Chromosome Abnormalities.* Cambridge: University of Cambridge Institute of Criminology. 1970. Pp. 117. £1.25.

WETTLER, M. *Syntaktische Faktoren im verbalen Lernen.* Bern: Huber. 1970. Pp. 151. D.M. 26.00.

WHITTAKER, J. O. *Introduction to Psychology* (Second edition.) Philadelphia: Saunders. 1970. Pp. xxii + 707. £4.57½.

WHITTAKER, J. O. *Student's Workbook to accompany 'Introduction to Psychology'.* (Second edition.) London: Saunders. 1970. Pp. 211. £1.60.

WILKINSON, R. *The Prevention of Drinking Problems: Alcohol Control and Cultural Influences.* London: Oxford University Press. 1970. Pp. xv + 301. £4.25.

WING, J. K. & BRANSBY, E. R. (eds.). *Psychiatric Case Registers.* London: H.M.S.O. 1970. Pp. 81. 90 np.

WOLSTENHOLME, G. E. W. & KNIGHT, J. (eds.). *Taste and Smell in Vertebrates: a CIBA Foundation Symposium.* London: Churchill. 1970. Pp. x + 402. £4.00.

WRENN, R. L. (ed.). *Basic Contributions to Psychology: Readings.* (Second edition.) Hemel Hempstead: Prentice-Hall. 1970. Pp. xix + 366. £2.50.

CHANGE-OVER TO SI UNITS

From January 1972 the journals of the British Psychological Society will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Further information about SI units is contained in the revised *Suggestions to Authors* pamphlet issued by the Society and obtainable, price 5*s*. (U.S.A. $1.00) post free, from Cambridge University Press.

Authors who refer to physical measures in their papers should now normally use SI units; common units of time (e.g. hour, year) will, of course, persist. Conversion tables will be published during the change-over period, but in the *British Journal of Psychology* these will be restricted to units of length, as others rarely occur. When they do, conversion factors will be stated in footnotes.

// REVUE DE PSYCHOLOGIE APPLIQUÉE

Publication Trimestrielle

Directeurs : Dr P. PICHOT et P. RENNES

3ème Trimestre 1970 – Volume 20 – No. 3

SOMMAIRE

Rédaction et Administration :
48 Avenue Victor Hugo, PARIS 16e

Abonnements (un an) France : 20 F – Étranger : 24 F
Le numéro France : 6 F – Étranger : 7 F

Réglements au nom des Éditions du Centre de Psychologie Appliquée

C.C.P. PARIS 5.157.61

Numéro spécimen sur demande

International Journal of Psychology
Journal International De Psychologie

Quarterly

Published by the International Union of Psychological Science

The *International Journal of Psychology* is devoted to cross-cultural comparisons of psychological phenomena in general psychology, child psychology and social psychology. It also deals with international confrontation of theories and methods on which psychologists from various countries differ because of their native culture and training. Oriented towards basic research and theory, it publishes scientific papers in English or in French. In each issue, an International Platform gives news and information on internationally significant problems.

Editor (Scientific papers) *G. de Montmollin*
28 rue Serpente Paris VI° France

Associate Editor (International Platform) *G. Nielsen*
University of Aarhus, Denmark

VOLUME 5 1970 No. 3

CONTENTS SOMMAIRE

Annual subscription (4 issues) $8.00 F 39 should be sent to DUNOD, 92 rue Bonaparte, Paris VI°

The British Journal of Psychology

The British Journal of Medical Psychology

The British Journal of Social and Clinical Psychology

The British Journal of Mathematical and Statistical Psychology

The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology, *The British Journal of Medical Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts, the subscription price, payable in advance, being £7.50 per volume post free ($25.000 in the U.S.A.) for *The British Journal of Psychology*, and £6.50 each ($22.00) for *The British Journal of Medical Psychology* and *The British Journal of Social and Clinical Psychology*. Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to Cambridge University Press, P.O. Box 92, London, N.W.1. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £6 per volume post free ($18 in the U.S.A.). Subscriptions should be sent to the Society's Agents, John Wright & Sons Ltd., the Stonebridge Press, Bath Road, Bristol BS4 5NU; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number. The subscription price in the U.S.A. is $4.20. Subscriptions should be sent to the Subscription Dept Manager, Associated Book Publishers Ltd, Distribution Centre, North Way, Andover, Hants.

The series of *Monograph Supplements* issued in connexion with *The British Journal of Psychology* is now discontinued. Orders may be sent to any bookseller, or Cambridge University Press, P.O. Box 92, London, N.W. 1 from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor W. Sluckin, Department of Psychology, The University, Leicester, LE1 7RH;

for *The British Journal of Medical Psychology* should be sent to

Dr J. J. Sandler, Room 101, Department of Psychiatry, Institute of Psychiatry, De Crespigny Park, London, S.E. 5;

for *The British Journal of Social and Clinical Psychology* should be sent to

Professor Thelma Veness (Social Psychology Editor), University of London Institute of Education, Malet Street, London, W.C. 1, *or* to
Dr J. G. Ingham (Clinical Psychology Editor), M.R.C. External Staff, Llandough Hospital, Penarth, Glamorgan, CF6 1XW;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Dr A. R. Jonckheere, Department of Psychology University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Professor J. D. Nisbet, Department of Education, University of Aberdeen, Old Aberdeen.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £8.50. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1.50 annually). *The British Journal of Medical Psychology* (£1.50 annually). *The British Journal of Social and Clinical Psychology* (£1.50 annually). *The British Journal of Mathematical and Statistical Psychology* (£1.50 annually), and *The British Journal of Educational Psychology* (£1.50 annually).

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18–19 Albermarle Street, London, W1X 4DN.

The British Journal of Psychology *Volume 62, Part 1, February 1971*

CONTENTS

Printed in Great Britain at the University Printing House, Cambridge

Volume 62, Part 2 *May 1971*

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by W. SLUCKIN

with J. ANNETT, W. S. ANTHONY, J. BROWN, R. CONRAD, N. F. DIXON, E. D. FRASER, J. A. GRAY, I. MARTIN, R. L. REID, H. R. SCHAFFER *and* P. B. WARR

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB

AMERICAN BRANCH:

32 EAST 57TH STREET, NEW YORK, N.Y. 10022

£2.50 net $8.50

Preparation of Papers for The British Journal of Psychology

In submitting papers for publication in *The British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this.
It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. 56 (1965) onwards.
A brief summary, suitable for use as an abstract, should precede the article.

Manuscripts should be submitted in duplicate.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation: and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.
Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.
The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.
Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–248.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *Année psychol.* **50**, 83–96.

World list abbreviations are used for journals.
The Oxford Dictionary is taken as the authority on spelling.

Subscriptions

The British Journal of Psychology is published quarterly.

Annual subscription (including postage) £7.50 net ($25.00 in the U.S.A.).

Second class postage paid at New York, N.Y.

Published by Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB and 32 East 57th Street, New York, N.Y. 10022.

Br. J. Psychol. (1971), 62, 2, pp. 147–150
Printed in Great Britain

THE EFFECT OF VOCALIZING ON COMPREHENSION IN THE PROFOUNDLY DEAF

By R. CONRAD*

Medical Research Council, Applied Psychology Unit, Cambridge

A group of hearing and a group of profoundly deaf school children were tested for comprehension after reading prose passages either silently or aloud. The deaf subjects were known, from previously published studies, to comprise a subgroup who primarily relied on articulatory coding to memorize verbal material, and another subgroup who seemed more to rely on a visual code. Neither the hearing controls nor the deaf articulators showed a significant effect of reading mode. The 'visualizers' comprehended significantly less when they read aloud than when they read silently. Though the two deaf groups performed equally well after silent reading, after reading aloud the comprehension difference was significant at better than the 0·001 level.

It is now well established that certain kinds of learning are facilitated when the material to be learned is heard or read aloud by the subject rather than read silently (Murray, 1965; Conrad & Hull, 1968; Murdock & Walker, 1969). However, in the reported cases it is recognized that the effect is limited to the most recently perceived material. For a number of authors (Craik, 1969; Crowder & Morton, 1969), this has been taken to support a multi-stage memorizing process, where for recent items acoustic information may be added to any other available. For the period of time when the acoustic information is still available, this provides an advantage over material which is only read silently.

Strong supporting evidence for this type of model was provided by Conrad (1970), who gave visually immediate serial recall tests of consonant sequences to profoundly deaf subjects. Conrad reported that analysis of errors indicated two clearly identifiable groups of subjects. One group evidently used an articulatory code in memory—confusing BCT and HX—but, unlike hearing subjects, showed no recall differences between reading the test sequences silently or aloud. This supports the memory models of Craik and of Crowder & Morton. A second group of deaf subjects did not confuse BCT or HX, but confused KXYZ. Conrad (1970) suggests that these subjects did not use articulatory coding but probably held the material in a visual code. This latter group recalled significantly worse when reading aloud than when reading silently. This observation has some relevance to the question of modality effects in short-term memory, but the study to be reported follows up a different implication.

The fact that approximately half the subjects tested were hindered in recall performance when obliged to read the test material aloud is relevant to widely used procedures in the education of deaf children. Where (as is common) there is strong emphasis on speech training, a great deal of ordinary school learning is accompanied by overt vocalization. It seems possible therefore that many children, who may or may not improve their speech thereby, might very well be handicapped in their learning. This was clearly so for immediate recall. The present study considers the case where more complex material is involved, namely prose passages read for comprehension.

* Present address: Nuffield Hearing and Speech Centre, 330 Grays Inn Road, London, W.C. 1.

METHOD

Subjects. Twenty-three profoundly deaf schoolboys who had been used in the study reported by Conrad (1970) were again used. On the basis of their memory errors each subject had previously been classed as either primarily using articulatory coding in short-term memory (A group) or primarily using visual coding (non-A group). There were 12 subjects in the A group (mean age: 13½ years), but only 11 were available to form the non-A group (mean age: 14½ years); this exhausted the school's population for which classificatory coding data had been obtained. The hearing loss of all subjects was such that it is safe to assume no auditory feedback when vocalizing at ordinary reading amplitude levels.

Twelve hearing children (mean age: 9½ years) were taken at random from an ordinary state school.

Materials. Six prose passages, each of about 300 words, were selected from a standard school collection of intermediate comprehension passages. Whilst a rough attempt was made to equate for difficulty, the experimental design was such that this was not a requirement. The passages were partly narrative and partly dialogue. For each passage, eight questions were prepared concerned with matter in the appropriate passage. Each question was in the form of four-item multiple-choice, with only one choice correct in every case. There was a separate question sheet for each passage. The form of the questions was such that the correct answer was not explicitly given (e.g. proper names). Above chance correct response could only be achieved through original comprehension of the meaning of the prose. The passages themselves were typed in double-spacing and pasted on to a card so that the final line was on the reverse side. This was an attempt to minimize rereading, but in fact the results indicate that this was not a relevant factor.

Procedure. The deaf subjects were tested either individually or in pairs. Even when reading aloud this was quite acceptable because of their absence of hearing. The general procedure was that the subject was given a passage, told to read it either silently or aloud, and as soon as he had finished was given the question sheet to complete. Exactly the same procedure was followed for the hearing subjects, with the proviso that subjects were individually tested.

The complete experimental design was such that 'silent' and 'loud' passages were alternated for each subject. Each of the six passages was read both silently and aloud equally often in the overall design as well as being completely counterbalanced for order of presentation. This design called for 12 subjects for completion. This was possible for the hearing subjects and the A group deaf subjects. As has been mentioned, only 11 non-A group deaf subjects were available, and the final row of the design was omitted. Though of course statistical purity was thereby lost, it seems unlikely that one more subject could have affected the empirical outcome to any serious extent. Since each subject was tested on every passage, he was tested on only half the passages reading silently, and half reading aloud. A complete test session took about 30 min. per subject, including a short preliminary separate practice passage which ensured that the instructions were fully understood.

The instructions were given in writing to the deaf subjects and verbally to the hearing subjects.

RESULTS

The measure used to compare reading silently with reading aloud was the percentage of wrong answers. Table 1 shows this for the hearing subjects, the deaf A group and the deaf non-A group. Although the hearing subjects were 4–5 years younger than the deaf, they showed much better comprehension overall. This particular comparison is irrelevant to the inquiry and the effect is in line with many published data (Furth, 1966). The main inquiry concerns the effect of reading mode.

When the distributions of errors as between the two modes are compared for the three subject groups, there is no difference between those for the hearing subjects and A group deaf, but the non-A deaf group is significantly different ($P < 0{\cdot}05$) from either. Where the former groups show improved comprehension when reading aloud, the non-A group deaf perform better when reading silently. In fact, the silent–loud

difference does not reach the 0·05 level for the hearing and A group deaf, but it is significant in the opposite direction for the non-A group deaf ($P < 0 \cdot 05$).

The most striking comparison perhaps derives from the quite fortuitous fact that when reading the passages silently, the difference in comprehension performance between the A group and the non-A group deaf subjects is not significant ($\chi^2 < 1$), i.e. the two groups are effectively matched. But when the reading mode is aloud, the difference, to the disadvantage of the non-A group, is significant at the 0·001 level ($\chi^2 = 19 \cdot 26$).

Table 1. *Percentage wrong answers*

Reading silently			Reading aloud		
Hearing subjects	A group deaf	Non-A group deaf	Hearing subjects	A group deaf	Non-A group deaf
41·6	58·0	61·0	37·2	51·4	69·7

Discussion

Unlike data from short-term memory studies, when prose is read for comprehension —and very much more material presented—there seems little advantage in overt vocalization for hearing subjects with this type of material. This agrees with Poulton & Brown (1967). For those deaf subjects who appear to use articulatory coding when committing verbal material to memory (A group), again overt vocalizing makes little difference to comprehension. This entirely agrees with Conrad's (1970) data, which showed the same effect for immediate recall. One can use the word 'agrees' here because the absence of effect in the short-term memory task was attributed to absence of auditory feedback. It would simply have been puzzling had the A group deaf comprehended more through reading aloud. The crucial result is the non-A group's performance. For these subjects reading aloud is clearly detrimental to comprehension of what they are reading. The reason for this might well be that suggested by Conrad for the detrimental effects of vocalizing in short-term memory; namely the distracting nature of vocalizing when the input processes for the material being learned are visual or analogues of visual. The fact that in silent reading there is no difference between the A group and non-A group is proof enough that the material can be learned. Vocalizing damages something somewhere along the learning line.

Impaired performance in some profoundly deaf subjects has now been shown on two widely different verbal learning tasks when these subjects are obliged to vocalize. Since at present we do not understand what it is that determines whether a particular deaf subject predominantly uses articulatory coding or 'visual' coding in these tasks, we have no *a priori* estimate of the relative proportions of these populations amongst the deaf. Were there evidence, either inferential or empirical, that the so-called non-A population were large, the results of these studies would cautiously contribute a qualification to universal oral instruction of the deaf that would be hard to ignore.

The deaf subjects were pupils at the Burwood Park School. The author is deeply indebted to the subjects, the Headmaster and the Board of Managers for their unstinted cooperation. The hearing subjects were juniors at the Normand Park Primary School and the cooperation of the school staff and of the Inner London Education Authority is gratefully acknowledged. Mrs B. C. Weiskrantz collaborated throughout.

References

Conrad, R. (1970). Short-term memory processes in the deaf. *Br. J. Psychol.* **61**, 179–195.

Conrad, R. & Hull, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychon. Sci.* **10**, 135–136.

Craik, F. I. M. (1969). Modality effects in short-term storage. *J. verb. Learn. verb. Behav.* **8**, 658–664.

Crowder, R. G. & Morton, J. (1969). Precategorical acoustic storage (PAS). *Percept. & Psychophys.* **5**, 363–373.

Furth, H. G. (1966). *Thinking Without Language.* London: Collier-Macmillan.

Murdock, B. B. & Walker, K. D. (1969). Modality effects in free recall. *J. verb. Learn. verb. Behav.* **8**, 665–676.

Murray, D. J. (1965). Vocalization-at-presentation and immediate recall, with varying presentation-rates. *Q. J. exp. Psychol.* **17**, 47–56.

Poulton, E. C. & Brown, C. H. (1967). Memory after reading aloud and reading silently. *Br. J. Psychol.* **58**, 219–222.

(*Manuscript received* 1 *April* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 151–156
Printed in Great Britain

DIFFERENTIAL EFFECTS OF IRRELEVANT DIMENSIONS IN THREE SHAPE RECOGNITION TASKS

By GILLIAN COHEN

Institute of Experimental Psychology, University of Oxford

Three shape recognition tasks are examined: matching two shapes, matching a written name to a shape, and naming shapes orally. In each the relevant dimension is shape, but the irrelevant dimensions of colour, orientation and proportion are varied. The effect of changes in the irrelevant dimensions in each task is used to infer the degree of specificity present in the internalized standard against which the test stimulus is matched. The fact that name-visual matches resemble visual-visual matches rather than naming latencies indicates that comparison in this task is based on a generated visual code.

Egeth (1966) found that subjects were unable to ignore irrelevant dimensions present in pairs of simultaneously presented multidimensional stimuli when making same–different judgements based on relevant dimensions. The present experiments examine the effect of irrelevant dimensions in three other types of matching task.

(1) In the first task the stimuli were presented successively; the first stimulus was the standard, and was held in short-term memory for comparison with the second, the test stimulus. The subjects were required to make same–different judgements based on shape alone. The stimuli were always either triangles, diamonds or rectangles, but the dimensions of colour, orientation and proportion were varied independently. Since the shapes remain invariant over transformations of these dimensions, they were always irrelevant to the judgement. The amount of interference, if any, caused by changes of an irrelevant dimension can be measured by comparing reaction times for identical pairs, with reaction times for pairs differing along this dimension.

(2) In the second task the subject was presented successively with the written name of a shape, followed by one of the three shapes, the irrelevant dimensions again being varied. Reaction times were compared for different values of each of the irrelevant dimensions to determine whether subjects formed consistent specific expectancies for these. In this type of verbal–visual matching task it is not easy to infer the subject's procedure. Clearly he must convert or transform one stimulus to a form equivalent to the other in order to make the match. Two strategies are open to him. (*a*) He can generate a visual coding of the verbal stimulus, and make a visual–visual match, as in (1), when the test stimulus appears. In an auditory–visual matching task using letters, Posner *et al.* (1969) concluded that subjects adopted this method, and Tversky (1969) has also demonstrated verbal-to-visual recoding. (*b*) He can retain the name of the first stimulus, name the test stimulus and make a verbal–verbal match.

(3) In the third type of task the subject was required to name the shapes which were visually presented. In this task the subject must match the stimulus against a standard which is a stored representation of the shape in long-term memory. The naming latencies were compared for different values of the irrelevant dimensions.

Logically the irrelevant dimensions might interfere with the subject's performance in the following ways: (*a*) The time to abstract the shape dimension from the test

stimulus, or to discriminate one shape from the other possible alternatives, might vary with the values of the irrelevant dimensions, e.g. it might be faster to identify a triangle when it is upright, and slower when it is tilted. Since the same set of test stimuli are used in all three tasks, interference at the stage of stimulus processing would affect all three tasks equally. (*b*) Interference from the irrelevant dimensions might also arise if the standard against which the test stimulus is matched contained irrelevant features, i.e. the representation stored in memory or generated is not an abstract one consisting of the shape alone, but a specific one; e.g. a triangle of a particular colour, proportion and orientation. Discrepancy between values of the irrelevant features of the standard and of the test stimulus would then impair performance at the comparison stage. Differences between the three tasks might exist if representations were more or less abstract, depending on whether they were in short-term memory, in long-term memory or generated.

Absence of interference from the irrelevant dimensions would only be found if the test stimulus could be identified with equal ease irrespective of the values of the irrelevant dimensions; and if the standard were either completely abstract, contained all possible values of the irrelevant dimensions simultaneously, or varied randomly between all the possible values.

The present experiment, then, examines the following questions: (1) Do changes in the irrelevant dimensions affect matching times? (2) If so, are the three types of task equally affected? If they are not equally affected, it may be inferred that this inequity arises from a difference in the degree of abstraction, or specificity in the stored representation of the shape. (3) If differences between the tasks are manifest, is the verbal–visual matching task most similar to the first, the visual–visual matching? Such a similarity would suggest strategy (*a*) with generation of a visual code facilitating a visual–visual comparison. Or is this task most similar to the third, the naming task, in which case strategy (*b*) would appear more likely.

The question of whether memory images are abstract or specific has long been a controversial one. Attneave (1967) echoed the view expressed by Berkeley in 1710 when he said: 'It is a curious fact that we cannot have a universal image... an image seems inexorably to demand values on some, though not necessarily all, of the dimensions over which object identification is invariant'. However, when the term 'image', suggesting, as it does, a kind of internal photograph, is discarded, abstraction is much less implausible. Hebb (1968) suggests that the hypercomplex cells identified by Hubel and Wiesel could form the basis of abstract memory representations of shapes. The present experiment is an attempt to submit the controversy to empirical tests.

METHOD

Materials. Each stimulus was one of three shapes: triangle, diamond or rectangle. Each shape had three irrelevant dimensions: colour, orientation and proportion. Colour was either red or blue; orientation was either tilted 45° to the left or upright; the proportion was varied by a constant amount such that the shape was either short and wide or tall and narrow. The shapes were cut from gummed coloured paper and mounted on plain white 6 × 4 in. cards.

Procedure. In the first two tasks the visual–visual match (VV) and the name–visual match (NV) the procedure was as follows. The subject viewed the cards through a tachistoscope. A blank field with a central fixation point was followed, after a verbal warning signal, by the first stimulus (S_1), which was exposed for 1 sec. After a fixed ISI of 750 msec. (offset to onset) the

second stimulus (S_2) was exposed. The subject's task was to judge it same or different, and to make a key-press response which terminated the exposure of S_2. The assignment of 'same' and 'different' response keys to right or left hands was balanced across subjects. The subject was instructed to ignore the irrelevant attributes of colour, orientation and proportion as far as possible, and to base his judgement on shape alone. He was also instructed to respond as fast as possible while trying to avoid errors. An Advance timer was started by the onset of S_2 and stopped by the key-press response. The stimuli subtended a visual angle of 2·7° approximately. The ISI of 750 msec. was selected since this was the interval at which Posner found auditory–visual matching equalled visual–visual matching of identicals; he concluded that generation was optimal at this interval.

Six male undergraduates served as subjects. They were paid volunteers, had no previous experience of similar experiments, and were unaware of the purposes of this experiment. All were right-handed. Reaction times and errors were recorded, and subjects were informed of both at each trial.

In the VV task S_1 consisted of a shape, and was followed by a second shape, S_2, to be judged same or different. In the NV task S_1 consisted of a typed word naming the shape ('triangle', 'diamond' or 'rectangle'), and was followed by a shape S_2, which was to be judged same if it corresponded to the name, and different if it did not. Thirty-six trials were same, and 36 were different. In the VV task the pairs which were same were divided into the following types. Group A: S_1 was identical with S_2 in all respects. Group B: S_1 was the same shape, colour and proportion as S_2 but differed in orientation. Group C: S_1 was the same shape, proportion and orientation as S_2, but differed in colour. Group D: S_1 was the same shape, colour and orientation as S_2, but differed in proportion.

The order of trials was randomized so that the various types of trial were randomly distributed throughout the series. In the NV condition the S_2 stimuli were exactly as in the VV task, and in the same order.

Half the subjects worked through the VV task first, and half had the NV task first. Subjects attended for two sessions, working through one task at each session. At the beginning of each session the subject had a block of practice trials of the type (NV or VV) assigned for that session. All combinations of shapes and irrelevant dimensions were equally represented in the practice trials, which were continued until the subject's performance stabilized, which was usually at about 30–40 trials. Subjects rested briefly halfway through each session. The intertrial interval was about 10 sec. and the sessions lasted about 50 min., and were separated by 3 days.

A third task (NL) to measure naming latencies was added subsequently. The S_2 stimuli used in the previous two tasks were exposed tachistoscopically. The onset of the stimulus triggered the timer. The subject was required to say the name of the shape, again ignoring the irrelevant dimensions. A throat microphone picked up the verbal response, and a voice key stopped the timer. Naming latencies were measured. Since it proved impossible to recall the subjects used in the VV and NV tasks, 12 new subjects were employed, the larger sample being necessary to offset individual differences.

Results

The analysis of the results is designed to answer the first two questions posed in the introduction to this paper.

(1) Does recognition time vary significantly with changes in the irrelevant dimensions?

The three tasks are analysed separately.

(*a*) *Visual–Visual matching.* Mean reaction times for judgements of 'same' were compared for group A (identicals): mean = 429 msec.; group B (orientation different): mean = 477 msec.; group C (colour different): mean = 457 msec.; and group D (proportion different): mean = 508 msec. The error rates for the four groups are fairly stable. Error percentages range from 1·8 to 3·9 per cent. A two-way analysis of variance, subjects × groups, yielded $F = 5{\cdot}23$; d.f. = 3, 15; $P < 0{\cdot}025$. The Newman–Keuls procedure was used to probe the differences between individual means. Groups

B and D both differ at the 0·01 level from group A, i.e. matching time is slower when either orientation or proportion is shifted, but not when colour is changed.

(*b*) *Name–Visual matching.* In this condition the name was matched against a shape which took one of two alternative values on three dimensions. There are therefore no identical pairs to serve as a base-line for comparison, and analysis was directed to the differences between values of each dimension in the judgements of same. On the null hypothesis there would be no significant differences between values of irrelevant dimensions. Inspection of the data reveals the existence of considerable differences. However, these differences are obscured in the analysis of variance by

Table 1. *Mean RTs (in msec.) for each subject's NV Same judgements, showing differences between the latencies associated with the two values of each irrelevant dimension*

	Colour		Orientation		Proportion	
Subject	Red	Blue	Upright	Tilted	Narrow	Wide
1	544	618*	524	626*	544	619*
2	567	512*	564	522*	551	534
3	561	488*	523	533	521	534
4	444	544*	436	568*	458	588*
5	329	368*	342	386*	328	382*
6	354	337	332	363*	343	351

* Significant differences. No subject had more than one error in any one category.

inconsistent trends, some subjects being faster for red shapes, for example, and some for blue. A three-way analysis of variance, subjects × dimensions × values of dimensions, yields a significant interaction of subjects × values of dimensions ($F = 32·99$; d.f. = 5, 10; $P < 0·01$). The presence of such a significant interaction will license the use of the Newman–Keuls test (Winer, 1962, p. 238) to test the differences between all possible pairs of means. Five subjects show differences significant at $P < 0·01$ for colour, five for orientation and three for proportion.

(*c*) *Naming latencies.* Again the analysis probed the differences between values of the irrelevant dimensions. In this task the differences are consistent, red being faster than blue, upright and narrow faster than tilted and wide. Mean reaction times pooled across subjects for judgements of 'same' are: red, 666 msec.; blue, 689 msec.; upright, 655 msec.; tilted, 694 msec.; narrow, 662 msec.; wide, 699 msec. Three separate analyses of variance were performed for colour, orientation and proportion. The colour factor ($F = 4·64$; d.f. = 1, 11) does not reach statistical significance. Orientation ($F = 12·06$; d.f. = 1, 11) and proportion ($F = 16·07$; d.f. = 1, 11) are both significant at $P < 0·01$.

(2) Do shifts in the irrelevant dimensions affect the three matching tasks differently?

To examine this question the mean difference between the latencies associated with the two values of each dimension, irrespective of the sign of the difference, was calculated for each matching task. Three Kruskal–Wallis one-way analyses of variance (Siegel, 1956, p. 184) were performed on the ranked differences under each dimension separately, to test the null hypothesis that the magnitude of the differences did not vary significantly. The Kruskal–Wallis procedure was selected for its capacity

to handle samples of unequal sizes. All the analyses yielded an H statistic significant at $P < 0{\cdot}001$ (colour, $H = 77{\cdot}8$; orientation, $H = 74{\cdot}1$; proportion, $H = 76{\cdot}7$). Individual comparisons between pairs of tasks were made by the Mann–Whitney U statistic (Winer, 1962, p. 623; Lewis & Cotton, 1958). Within the colour dimension, NL and VV both differ from NV with U values significant at $P < 0{\cdot}05$. Within the orientation dimension, NL differs from both VV and NV at $P < 0{\cdot}05$, and within the proportion dimension NL differs from VV, again at $P < 0{\cdot}05$.

Table 2. *Mean differences between the latencies associated with the two values of each dimension (in msec.) for each of the three tasks*

Tasks	Colour	Orientation	Proportion
VV	29·6	54·5	64·0
NV	59·6	60·0	49·5
NL	33·2	44·2	43·1

Discussion

The effects of changes in the irrelevant dimensions. Shifts of colour do not significantly affect the visual–visual match or the naming latencies. It follows that the shape can be discriminated equally well whether red or blue. Where an irrelevant dimension causes no interference it is not possible to conclude that the stored representation against which the stimulus is matched, is abstract with respect to this dimension. The result would occur if the stored representation contained all values of that dimension, or if it shifted randomly between alternative values. The NV task differs from the other two in that five out of six subjects had significant biases towards one of the colour values. All three tasks show significant effects of shifts in the dimensions of orientation and proportion.

Are the tasks differentially affected by the irrelevant dimensions? Section 2 of the results confirms that such a difference exists. The naming latencies are less affected than the NV task by colour and by orientation values. The VV task differs from the NV task in colour, but not in the other two dimensions. These differences cannot arise at the stimulus processing stage, as the same test stimuli are used in each task and the three tasks are therefore equivalent in this respect, and must therefore arise at the stage of comparison of the test stimulus with an internalized representation of the standard. Where matching times are significantly faster for one particular value of a dimension, this must be because the standard contains this value and comparison is faster when the test stimulus coincides.

Is the NV task more similar to the VV or to the NL task in its sensitivity to shifts in irrelevant dimensions? It appears to be marginally more similar to VV matching, since it differs from this on only one dimension, while it differs from the NL task on two dimensions. While this partial resemblance is hardly sufficient grounds for concluding that subjects in this task generate a visual coding and perform a visual match rather than naming the test stimulus and performing a name match, some additional evidence supports such a hypothesis. First, asked informally to describe their procedure in the NV task, four out of six subjects reported that they formed an image which had specific values on all dimensions including colour. Secondly, while subjects

in the NL task all showed a bias towards the *same* value of each dimension, subjects in the NV task varied, so that, for example, some were faster when the test stimulus was red, and some when it was blue. The lack of uniformity in the biases displayed in this task may reflect a more temporary and labile internalized standard.

Comparison of overall mean reaction times will not serve in this experiment to determine the strategy employed in the NV task. Following the argument of Posner *et al* (1969) the mean of the NV same scores should equal the mean of the VV identicals if a visual representation is being generated from the name. However, this line of argument will not hold where the stimuli are multidimensional. In the VV identical group the coincidence of the irrelevant attributes would probably facilitate judgement. If the generated representation were specific, it could not have the correct values of all the attributes at every trial, and if it were abstract it would still lack the facilitating effect of coinciding irrelevant attributes. The fact that the NV same mean (476 msec.) exceeds the VV identical mean (430 msec.) cannot be taken as evidence that visual matching does not occur in the NV task. It does not correspond extremely closely to the VV same mean (all groups included) of 474 msec., and this supports the conclusion that the same matching process is in operation in both tasks.

In conclusion, it can be stated that all forms of recognition examined here are to some extent affected by the presence of irrelevant attributes; that an irrelevant colour is more easily excluded from comparison; that matching against a permanent long-term standard, as in shape-naming, is relatively less affected than matching against a temporary short-term standard. Although the evidence must be regarded as suggestive rather than conclusive, cross-modal verbal–visual matching probably takes place by converting the verbal label to a visual code, rather than vice versa.

This work was carried out while the author was in receipt of a Medical Research Council scholarship.

REFERENCES

ATTNEAVE, F. (1967). Criteria for a tenable theory of form perception. In W. Wathen-Dunn (ed.). *Models for the Perception of Speech and Visual Form*. Cambridge, Mass.: M.I.T. Press.

BERKELEY, G. (1710). *The Principles of Human Knowledge*. Oxford: Clarendon Press, 1901.

EGETH, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Percept. & Psychophys.* **1**, 245–252.

HEBB, D. O. (1968). Concerning imagery. *Psychol. Rev.* **75**, 466–477.

LEWIS, D. J. & COTTON, J. W. (1958). Partial reinforcement and nonresponse acquisition. *J. comp. physiol. Psychol.* **51**, 251–254.

POSNER, M. I., BOIES, S. J., EICHELMAN, W. H. & TAYLOR, R. L. (1969). Retention of visual and name codes of single letters. *J. exp. Psychol. Monogr.* **79**, no. 1, pt. 2, 1–16.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.

TVERSKY, B. (1969). Pictorial and verbal encoding in a STM task. *Percept. & Psychophys.* **6**, 225–233.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

(*Manuscript received* 8 *March* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 157–163
Printed in Great Britain

THE INFLUENCE OF FORCE MAGNITUDE ON THE PERCEPTION OF BODY POSITION

I. EFFECT OF HEAD POSTURE

BY N. J. WADE* AND H. SCHÖNE

Max-Planck-Institut fur Verhaltensphysiologie, Seewiesen, West Germany

Judgements of body position during rotation in the sagittal plane were investigated at force magnitudes of 1·0, 1·2, 1·4, 1·6 and 1·8 *g*. With the head and trunk upright (Expt. I) the judgements were displaced forward as a function of the force operating. The forward shift could be related to the increase in the backward acting shear force on the utricular statoliths. This hypothesis was supported by the reduced forward shift in perceived body position that occurred under the same conditions but with the head tilted 30° forward of the upright trunk (Expt. II).

During passive rotation the upright position of the body can be judged accurately without the aid of visual cues (see Howard & Templeton, 1966). Considering that the vestibular system plays an important role in the maintenance of the normal upright posture it would also be expected to be involved in this task. However, the evidence appears contrary to this notion: subjects without functional vestibular systems locate the postural upright with almost the same accuracy as normal subjects (Clark & Graybiel, 1963; Garten, 1920), although the former yield a larger postural after-effect following prolonged tilt than the latter (Clark & Graybiel, 1964). It is, of course, possible that the vestibular defective subjects have learned to utilize the information from their remaining postural systems in a manner similar to the use of vestibular information in normal subjects.

Much of the research in this area has been concerned, somewhat unsuccessfully, with unravelling the methodological complexities involved in determining the postural upright. The main problem is that the system under investigation is undergoing constant stimulation, i.e. the postural receptors are always stimulated by the force of gravity. This has limited the psychophysical procedures employed almost entirely to the method of adjustment. With this method the body is initially rotated to a non-upright position and then returned towards the upright. In reaching the non-upright (starting) position sufficient information has already been given to the subject to allow him to make an adjustment with some accuracy, using only the time cue. This possibility is increased when the body is returned to the upright position after each adjustment, as is usually the case. Add to this the experiments that have used only one starting position (but generally two directions) and it is conceivable that the time cue may be the major source of information utilized. These factors may have influenced the performance of the vestibular defective subjects in the studies cited above. It is not suggested that subjects deliberately use the time cue in making an adjustment, rather that this is always a factor in these experiments and cannot be excluded from consideration.

The experiments reported herein were designed to examine whether differential stimulation of the statolith (or otolith†) system results in variation of perceived body

* Present address: Department of Psychology, University of Dundee.

† The terms statolith and otolith tend to be used interchangeably, but the former is recommended in *Nomina Anatomica* (1966).

position. The statolith organs consist of the paired utricles and saccules, and attention is directed to the former. When the head is held so that the line joining the external auditory meatus and the base of the orbit is horizontal the utricular maculae are inclined posteriorly by about 28° (Corvera *et al.*, 1958). The effective stimulus for the statolith organs is the shear force, which acts tangentially to the macular surface (von Holst, 1950; Jongkees, 1969; Trincker, 1962). In a body sagittal plane the utricular shear force may be increased either by tilting the head backwards or by increasing the force magnitude operating. The latter procedure was employed here, using a centrifuge. It was hypothesized that an increase in the backward acting shear force would result in forward inclination of the body in order to obtain the equivalent perceived body position to that under normal gravity. It has been demonstrated that force changes produced either by oscillation on a parallel swing or during centrifugation with the subject facing the centre of rotation lead to an apparent tilting of the body (Graybiel, 1956; Jongkees, 1952). Under similar conditions to those employed here the subjective visual horizon varies as a function of the utricular shear force, in addition to changes in apparent body position (Schöne, 1962, 1964).

Method

The method employed in the present experiments differed somewhat from those used in previous studies. First, the subject was rotated in the sagittal plane. Most experiments have involved rotation in the lateral plane, and the postural upright can be located more accurately in this than in the sagittal plane (Burtt, 1918; Clegg & Dunfield, 1954*a*, *b*; Kleinknecht, 1922; Passey & Guedry, 1949). Secondly, the subjects were trained to judge the upright body position accurately under normal gravity. Errors in adjusting the body to the upright are reduced by practice with feedback (Clark & Graybiel, 1963; Fleishman, 1953; Garten, 1920; Pearson & Hauty, 1959, 1960; Solley, 1956, 1960), and this fact was exploited to achieve a certain criterion of performance before commencement of the experimental series. This was done so that the subjects would have a clear impression of the stimulus complex defining 'upright', and that they would be able to equate for this at increased force magnitudes. Moreover, subjects received additional feedback trials under normal gravity before each session. Thirdly, within a session the subject was not returned to the trained upright position after each adjustment, and the subsequent starting position was determined from the position of the previous judgement.

Subjects. Six subjects participated in both experiments; none of them had any known deficits in vestibular function.

Apparatus. The body supports for the seated subject consisted of an individual dental composition bite-board and adjustable chest and leg supports, in addition to the back, seat, and foot rests of the chair. The chair rotated in a sagittal plane, around an axis that approximated the line joining the labyrinths, at a velocity of 2°/sec. Its orientation, which could be read to the nearest 0·5°, was controlled by the experimenter. The chair was situated in a darkened gondola which hung from the end of an arm of a large centrifuge. The gondola was free-swinging and always assumed the direction of the resultant force during centrifugation, so that only the magnitude of the force operating on the subject was increased. The force strength was registered within the gondola and could be read by the experimenter to the nearest 0·01 g. Verbal communication between the subject in the gondola and the experimenter in the control room was possible at all times. Further details of the apparatus are given by Schöne (1962, 1964).

Procedure. The task involved the subject making a verbal signal when he felt that his body was upright. Before commencement of the experimental series the subject underwent a training session. This involved, first, six trials in which the chair was tilted forward (+) and backward (−), and returned by the experimenter exactly to the upright; secondly, a series of trials in which the subject indicated when his body felt upright. The starting positions of the chair were 10, 20, and 30° forward and backward, presented in random order. The chair was stopped after each response and the error was noted. The subject was returned to the upright and informed

that he was in the correct position. This procedure was continued until eight successive trials were completed with errors of 4° or less. The training series was given in the standing centrifuge.

The force magnitudes investigated were 1·0, 1·2, 1·4, 1·6 and 1·8 g. A session consisted of four practice trials at 1·0 g, similar to the first part of the training series, after which the centrifuge was accelerated slowly to reach the appropriate force magnitude. Six adjustments to the trained body position were then made from the above, counterbalanced starting positions, presented in random order. The chair was not returned to the correct position after each adjustment but remained in the position the subject indicated felt upright. The subsequent starting position was determined from this orientation. For example, if the subject felt upright with the chair tilted 5° forward and the next starting position was 20° forward the chair was rotated to 25° forward of the trained upright body position for the commencement of the next trial.

Each subject completed four sessions at each force magnitude. The order of the first five sessions, one for each force value, was randomized and this was successively reversed for each of the remaining three blocks of five sessions, so minimizing any order effects.

Experiment I

In this experiment the upright position of the body was defined as an approximately vertical torso position (corresponding to a specific chair orientation) with the teeth firmly impressed in the horizontal bite-board.

Results

The judgements of the body upright position were averaged over starting positions and sessions for each force magnitude, and the values are represented graphically in Fig. 1. The zero point on the ordinate represents the trained upright position. The curve has a significant linear trend component ($F = 79{\cdot}38$; d.f. $= 1, 20$; $P < 0{\cdot}001$). The residual trend was not significant. With increased force magnitude the perceived upright position of the body was displaced in the forward direction as predicted from the increased backward acting shear on the utricular maculae.

Experiment II

The results of the above experiment are strongly suggestive of statolith organ involvement in the perception of body position. However, it could equally well be argued that the shift in judgements is a function of the mechanoreceptors in the trunk, as these are similarly stimulated by the force operating. Therefore a more specific test for the involvement of the statolith system was designed. When the head is tilted forward the plane of the utricular maculae subtends a smaller angle to the horizontal than with the head upright. Accordingly, the increment of the utricular shear force consequent on increasing the force operating would also be smaller. On the other hand, with the trunk in the same upright position as in Expt. I its pressure receptors would be stimulated in the same way. According to the utricular shear hypothesis a smaller shift in judgements would occur with the head tilted forward.

The method employed in this experiment was the same as in Expt. I. The chair orientation for the trained upright position was the same as the above, but the head was tilted forward by 30° and controlled in that position by means of the bite-board.

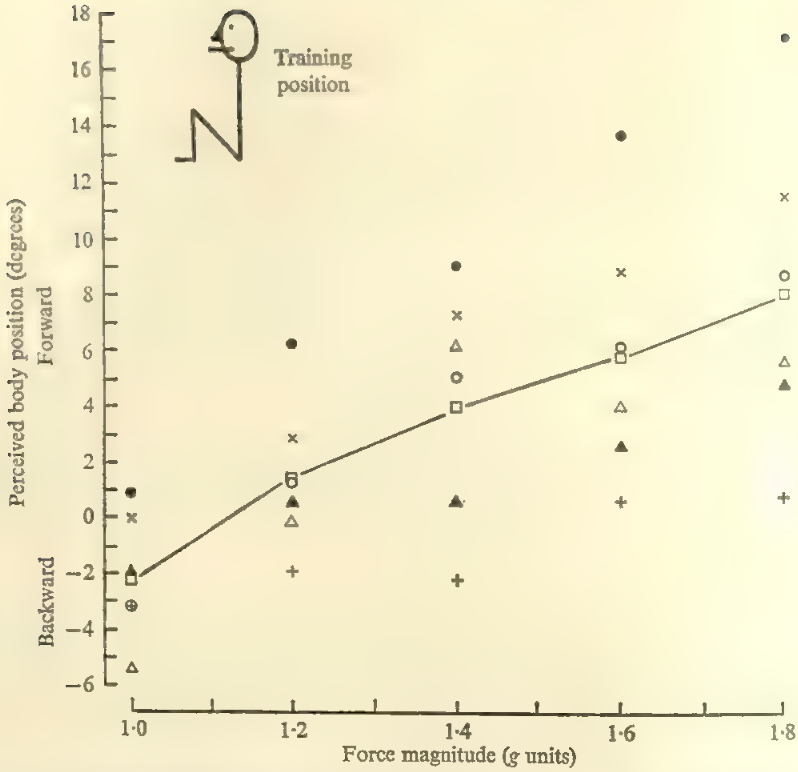


Fig. 1. Deviations of perceived body position from the training position as a function of force magnitude with head upright (Expt. I). The 0 point on the ordinate denotes the training position. The symbols represent the averaged data for each subject and the line joins the mean values.

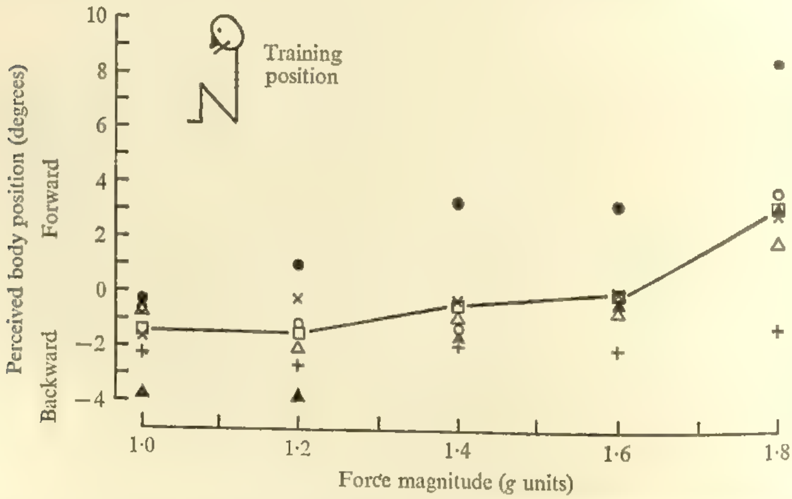


Fig. 2. Deviations of perceived body position from the training position as a function of force magnitude with head tilted 30° forward. The 0 point on the ordinate denotes the training position and the symbols represent the data from the same subjects as in Fig. 1.

Results

The judgements of body position were averaged over starting positions and sessions for each force magnitude, and the values are shown graphically in Fig. 2. The curve has a significant linear trend component ($F = 55{\cdot}32$; d.f. = 1, 20; $P < 0{\cdot}001$), and the residual trend also reached significance ($F = 4{\cdot}73$; d.f. = 3, 20; $P < 0{\cdot}05$). The perceived body position was shifted forward with increased force magnitudes, but to a much smaller degree than with the head upright.

Discussion

The hypothesis that the perceived body position is influenced by stimulation of the statolith system was supported by the above experiments. With increased force magnitudes the forward displacement of the body was greater when the increment in the backward acting utricular shear was greater. In a similar manner the visual horizon varies as a function of the shear force operating (Schöne, 1962, 1964). Nonetheless, the results seem to be at variance with the investigations of the comparative performance of normal and vestibular defective subjects (Clark & Graybiel, 1963; Garten, 1920). However, as mentioned above, it is possible that methodological factors permitting the utilization of non-postural cues may have acted to reduce any differences between the two groups. There is also the difficulty of assessing the degree of compensation that has developed in the vestibular defective subjects. The data here presented concerned the performance of normal subjects under a wide range of stimulation conditions, and did not involve problems associated with comparing subjects with and without some specific sensory deficit.

Although the subjects were trained to the same criterion before commencement of the experimental series there were still wide individual differences, even at $1{\cdot}0\,g$. In fact, in both experiments the body was tilted backward under normal gravity in order to be perceived as the trained position. In Expt. I all subjects showed the forward shift in judgements as a function of force strength, but to largely differing degrees, as can be seen in Fig. 1. Both the shift and the individual differences were smaller in Expt. II (see Fig. 2).

The results suggest that the perceived upright position of the body corresponds to a specific shear force acting on the utricular statoliths, and that this is equated at increased force magnitudes by forward inclination of the head. The shear force is computed as the product of the force magnitude and the sine of the angle of inclination of the maculae to the horizontal. With the bite-board horizontal (Expt. I) the line joining the external auditory meatus to the base of the orbit was tilted backwards such that the utricular maculae were inclined by approximately 40° to the horizontal. Accordingly, with the bite-board tilted 30° forward (Expt. II) the inclination of the maculae was about 10°. Taking these values into account it is clear that the equation of perceived body position and shear force is not exact. For instance, under normal gravity the backward acting utricular shear with the head upright (bite-board horizontal) was $0{\cdot}64\,g$; in order to obtain this shear at $1{\cdot}8\,g$ the head would need to be tilted forward by 19°, whereas the experimental value was 10°. The equivalent computation for the head tilted 30° forward is a 4° forward shift,

which corresponds to the obtained value; it should be noted, however, that the shift at 1·8 g was relatively greater than those at lower force magnitudes with the head so positioned (see Fig. 2). It is clear that either the relation between utricular shear and perceived body position is non-linear or other factors are involved in the judgements. Concerning the latter, certain aspects of the methodology may have resulted in a decreased forward shift in the perceived body upright. For example, the body was in the upright position during acceleration of the centrifuge so that the first starting position was determined from this orientation. This may have biased judgements in the direction of the trained upright position. Furthermore, although the subject was not returned to the upright after each adjustment the rotation time to the starting position could have been equated with the adjustment time. This factor, in combination with the previous one, may also have reduced any displacement of perceived body position. It is unlikely that the prolonged stimulation during centrifugation had a differential effect on the judgements because in a preliminary study the size of a postural after-effect following several minutes forward or backward tilt was not dependent upon the force magnitude.

The difference between the results of the two experiments cannot be attributed to the operation of the trunk system as this was in about the same orientation throughout. Nonetheless, there may have been a slight contribution from the trunk receptors in both experiments, either from internal pressure receptors or from cutaneous pressure differences. With respect to the latter, the pressure differences on the teeth in the bite-board may have influenced the judgements in the different head positions.

Little is known about the role of the saccular statoliths in the maintenance of posture. They would have been differentially stimulated by the increase in the force operating in both experiments, and their involvement in the changes that occurred cannot be discounted.

References

BURTT, H. E. (1918). The perception of slight changes of equilibrium with especial reference to problems of aviation. *J. appl. Psychol.* **2**, 101–115.

CLARK, B. & GRAYBIEL, A. (1963). Perception of the postural vertical in normals and subjects with labyrinthine defects. *J. exp. Psychol.* **65**, 490–494.

CLARK, B. & GRAYBIEL, A. (1964). Perception of the postural vertical following prolonged bodily tilt in normals and subjects with labyrinthine defects. *Acta oto-lar.* **58**, 143–148.

CLEGG, W. C. & DUNFIELD, N. M. (1954*a*). Non-visual perception of the postural vertical. I. Sagittal plane. *Can. J. Psychol.* **8**, 1–8.

CLEGG, W. C. & DUNFIELD, N. M. (1954*b*). Non-visual perception of the postural vertical. II. Lateral plane. *Can. J. Psychol.* **8**, 80–86.

CORVERA, J., HALLPIKE, C. S. & SCHUSTER, E. H. J. (1958). A new method for the anatomical reconstruction of the human macular planes. *Acta oto-lar.* **49**, 4–16.

FLEISHMAN, E. A. (1953). Perception of body position in the absence of visual cues. *J. exp. Psychol.* **46**, 261–270.

GARTEN, S. (1920). Über die Grundlagen unserer Orientierung im Raum. *Abh. sächs. Ges. (Akad.) Wiss.* **36**, 431–510.

GRAYBIEL, A. (1956). The importance of the otolith organs in man based upon a specific test for utricular function. *Ann. Otol.* **65**, 470–487.

HOLST, E. VON (1950). Die Arbeitsweise des Statolithenapparates bei Fischen. *Z. vergl. Physiol.* **32**, 60–120.

HOWARD, I. P. & TEMPLETON, W. B. (1966). *Human Spatial Orientation.* London: Wiley.

JONGKEES, L. B. W. (1952). Some remarks on the functions of the vestibular organs. *Rep. Inst. Lar. Otol.* **2**, 1–10.

JONGKEES, L. B. W. (1969). Physiologie und Pathophysiologie des Vestibularorganes. *Arch. klin. exp. Ohr.-, Nas.- u. Kehlk. Heilk.* **194**, 1–110.

KLEINKNECHT, F. (1922). Ein weiterer Beitrag zur Frage des Übungseinflusses und der Übungsfestigkeit am Neigungsstuhl. *Z. Biol.* **77**, 11–28.

NOMINA ANATOMICA (1966). Amsterdam: Excerpta Medica Foundation. 3rd ed.

PASSEY, G. E. & GUEDRY, F. E. (1949). The perception of the vertical. II. Adaptation effects in four planes. *J. exp. Psychol.* **39**, 700–707.

PEARSON, R. G. & HAUTY, G. T. (1959). Adaptive processes determining proprioceptive perception of verticality. *J. exp. Psychol.* **57**, 367–371.

PEARSON, R. G. & HAUTY, G. T. (1960). Role of postural experience in proprioceptive perception of verticality. *J. exp. Psychol.* **59**, 425–428.

SCHÖNE, H. (1962). Über den Einfluss der Schwerkraft auf die Augenrollung und auf die Wahrnehmung der Lage im Raum. *Z. vergl. Physiol.* **46**, 57–87.

SCHÖNE, H. (1964). On the role of gravity in human spatial orientation. *Aerospace Med.* **35**, 764–772.

SOLLEY, C. M. (1956). Reduction of error with practice in perception of the postural vertical. *J. exp. Psychol.* **52**, 329–337.

SOLLEY, C. M. (1960). Influence of head tilt, body tilt, and practice on reduction of error in perception of the postural vertical. *J. gen. Psychol.* **62**, 69–74.

TRINCKER, D. (1962). The transformation of mechanical stimulus into nervous excitation by the labyrinthine receptors. In *Biological Receptor Mechanisms: Symp. Soc. exp. Biol.*, no. 16. Cambridge: Cambridge University Press.

(*Manuscript received* 1 *April* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 165–173
Printed in Great Britain

THE EFFECT OF VARYING THE TIME INTERVAL BETWEEN EQUAL AND OPPOSITE CORIOLIS ACCELERATIONS

By JAMES T. REASON* and ASHTON GRAYBIEL
Naval Aerospace Medical Institute, Pensacola, Florida, U.S.A.

The purpose of this experiment was to investigate the effect of varying the time interval between two equal and opposite Coriolis accelerations upon the duration of the subjective responses evoked by the second stimulus. It was also designed to evaluate certain predictions generated from a 'torsion pendulum' model of the neural events mediating these subjective phenomena.

Theoretical curves derived from the torsion pendulum model approximated fairly closely the way in which the reported durations of the subjective phenomena increased as a function of the time interval between the two Coriolis accelerations. This result supported the *a priori* assumption that the neural events underlying the subjective phenomena are closely linked to mechanical events occurring within the cupula-endolymph system. However, an explanation resting entirely upon peripheral phenomena would be inadequate to account for two additional findings: (1) the estimated time constants of signal decay were shorter than those expected on the basis of the known mechanics of the semicircular canal system, and (2) the persistence of the Coriolis sensation (feelings of apparent whole body motion without visual reference) was greater at all intervals than the Coriolis oculogyral illusion. Adequate explanation of these findings requires the postulation of additional central mechanisms.

The Coriolis vestibular reaction occurs when a subject, seated on a platform rotating at constant velocity, moves his head about some axis other than the axis of platform rotation. During and immediately following the head motion, the subject experiences a sensation of apparent rotation about an axis which is orthogonal to both head-tilt and platform axes. Semicircular canals, previously unstimulated by the platform rotation, are brought into the plane of this rotation by the head motion, causing an accelerative stimulus to be delivered to the cupula-endolymph systems of these canals. When the head motion is of short duration, the strength of the Coriolis accelerative stimulus is determined by the angular velocity of the platform and the angle through which the head is moved. A good description of the mechanics of the Coriolis reaction is given by Guedry & Montague (1961).

The Coriolis vestibular reaction represents an interesting case of 'sensory rearrangement' in which the two vestibular receptors—the semicircular canals and the otoliths—simultaneously signal contradictory information concerning the orientation of the head. It is also extremely effective in provoking the symptoms of motion sickness. Providing he possesses an intact vestibular system, almost any individual can be made motion sick given sufficient quantity and duration of Coriolis stimulation. As a consequence of these effects, the Coriolis reaction is foreseen as a major hazard in projected spacecraft that rotate to provide artificial gravity.

This investigation was concerned with the situation in which a Coriolis accelerative stimulus (S_1) is followed after a variable time interval (I) by a second stimulus (S_2) that is equal in magnitude but opposite in direction to the first. The Coriolis stimulation was delivered to the semicircular canal system by a 45° head tilt to the left shoulder (S_1) followed, after the variable interval, by a return motion to the upright

* Present address: Department of Psychology, University of Leicester.

position S_2) during rotation at constant velocity. The effect of varying the time interval (I) was studied in relation to two measures of subjective response: (1) the duration of the sensation of pitching about the body's y-axis (termed the 'Coriolis sensation') and (2) the duration of the apparent motion of a dimly illuminated target (termed the 'Coriolis oculogyral illusion' or OGI).

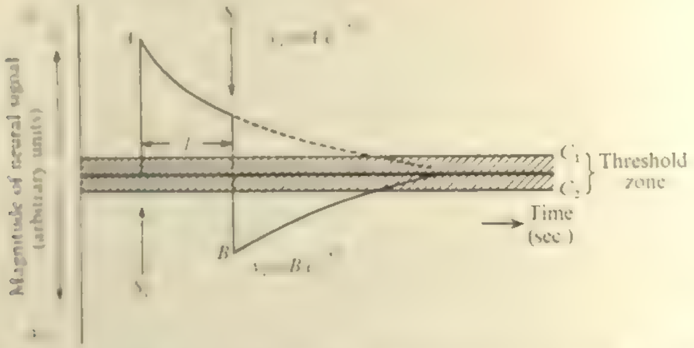


Fig. 1. Diagrammatic representation of the 'torsion pendulum' model. (See text.)

A 'torsion pendulum' model (see Fig. 1) was used to generate the experimental predictions. The basic assumption of the model was that the neural events underlying the perception of the subjective phenomena would be dependent upon the physical events occurring within the vestibular end-organ. This formulation allowed the known mechanical properties of the cupula-endolymph system (Groen, 1957; Van Egmond *et al.*, 1949) to be used as the basis for predictions concerning the duration of the subjective phenomena evoked by S_2.

The uppermost plot in Fig. 1 indicates the initial magnitude and time-course of the neural signal evoked by the first Coriolis acceleration, S_1. This neural activity is assumed to decay exponentially with a time constant of $1/k$ and has the form: $y_1 = Ae^{-kt'}$ (where y_1 = magnitude of the neural signal evoked by S_1; t' = time from onset of S_1). After 1 sec. an equal yet directionally opposed stimulus, S_2, is delivered to the cupula-endolymph system which causes the signal to reverse its sign in the manner shown in Fig. 1. The initial extent of this signal reversal due to S_2 is assumed to be equal to the peak signal value, A, evoked by S_1. The peak magnitude of the signal produced by S_2 is designated B, where $B = -A - Ae^{-kt'}$. This signal also decays exponentially with a time constant of $1/k$. Hence the neural response to S_2 is of the form: $y_2 = [A(1-e^{-kI})]e^{-kt}$ (where y_2 = magnitude of the neural response evoked by S_2; t = time after S_2). It is also postulated that the subjective phenomena mediated by y_2 will cease when $y_2 = C$, a threshold value. Thus the persistence of the subjective phenomena evoked by S_2 is expected to increase as a function of I up to a certain asymptotic value, as shown in the following equation:

$$T_{R_2} = (1/k)\log_e\left(\frac{1-e^{-kI}}{D}\right),$$

where T_{R_2} = duration of the subjective response evoked by S_2 (sec.), $1/k$ = time

constant of signal decay (sec.). $D = C/A$ or the ratio of peak to threshold signal, and I = interval between S_1 and S_2 (sec.).

Method

Subjects. Twenty young men, either Navy enlisted men or college students, served as subjects. All were known to possess normal vestibular function and to be in good health at the time of the experiment.

Procedure. The experiment was performed in the Pensacola Slow Rotation Room (SRR). A full description of this facility can be found elsewhere (Graybiel *et al.* 1960).

The subjects were seated with their backs to the inside wall of the SRR at a distance of 7 ft from the centre of rotation. The subjects' z-axes were aligned parallel to the axis of rotation. Head pads attached to the wall restricted the lateral head movements to a tilt through an arc of 45° to the left shoulder and back to the upright. The head movements were made at a uniformly rapid rate and were completed within 0·50–1·0 sec. The movements were practised while the device was stationary to establish a constant rate as far as possible. The rotation was in the counterclockwise (CCW) direction throughout.

In 13 subjects, the effects of the Coriolis accelerations were measured by recording the persistence of the OGI. The OGI target consisted of a dimly illuminated 6-in. cube placed at a distance of 10 ft. from the subject. Only the outline of the cube was illuminated, and it was orientated such that nine edges were visible to the subject. For these measurements, the SRR was in complete darkness except for the target and a small light attached to the experimenter's clipboard. The latter was screened from the subject by a curtain.

In the other seven subjects, the effects of the same accelerations were measured by recording the persistence of the Coriolis sensation, i.e. the time elapsing between the completion of the head movement and the point at which the subject reported no further sensation of apparent bodily motion. For these measurements, the subjects were blindfolded throughout the session.

The stimulus procedures were identical for both the OGI and sensation groups. Rotation started at 7·5 r.p.m. At this speed, each subject began by executing three sets of controlled tilt and return movements, where the interval between any two motions was kept fixed at 30 sec. On completion of each discrete movement, the subject indicated the persistence of his apparent visual or bodily motion. This procedure (termed 'fixed-interval motions') was included to provide a baseline persistence value for the subjective response to S_2 when the interstimulus interval was such that the response to S_1 would have largely dissipated before S_2 was applied. It also provided the subjects with an opportunity to stabilize their endpoint decision criteria prior to the variable-interval determinations.

Some 5 min. after the completion of the fixed-interval measurements, the variable-interval motions were commenced. The intervals (I) between S_1 (tilt) and S_2 (return) were: 1, 2, 4, 8, 16 and 32 sec. The movements were made on verbal instructions from the experimenter, and only the response to S_2 was recorded. The intervals were presented in a random order, and a different random sequence was used for each subject.

On completing these measurements at 7·5 r.p.m., the subjects were rested for approximately 5 min. with head immobilized. The angular velocity of the SRR was then increased to 10 r.p.m. where fixed- and variable-interval movements, identical to those described above, were executed.

One departure from this procedure was the fact that the 32-sec. interval was not included in the design until five of the OGI subjects had been tested. Thus, while persistence values for this interval were obtained for all seven of the 'sensation' subjects, they were obtained for only eight of the OGI subjects.

Results

The mean persistence values evoked by S_1 and S_2 during the initial fixed-interval movements at 7·5 and 10 r.p.m. are shown in Tables 1*a*, *b*. Table 1*a* shows the duration values for the OGI subjects, while Table 1*b* shows the durations for the sensation subjects. Grand means and standard deviations are given at the bottom of each table.

Inspection of these tables indicates that S_2 evoked a more persistent response, both for OGI and sensation subjects, than the tilt motion S_1. The Wilcoxon matched-pairs signed ranks test (Siegel, 1956) showed that these differences were significant for all comparisons: OGI at 7·5 r.p.m. ($T = 6{\cdot}5$; $P < 0{\cdot}01$, two-tailed test); OGI at 10 r.p.m. ($T = 2{\cdot}0$; $P < 0{\cdot}01$, two-tailed test); Coriolis sensation at 7·5 r.p.m. ($T = 0$; $P < 0{\cdot}05$, two-tailed test). Coriolis sensation at 10 r.p.m. ($T = 0$; $P < 0{\cdot}02$, two-tailed test).

Table 1*a*. *Mean persistence of the OGI following three sets of fixed-interval (30 sec.) 45° tilt and return motions*

	Counterclockwise rotation			
	7·5 r.p.m.		10 r.p.m.	
Subject	Tilt (S_1)	Return (S_2)	Tilt (S_1)	Return (S_2)
1	3·2	3·5	2·2	2·3
2	3·0	3·7	2·3	4·8
3	10·0	11·0	7·8	9·5
4	1·0	1·2	1·0	1·3
5	5·2	8·5	3·5	8·0
6	4·2	5·0	3·3	3·8
7	1·8	2·0	1·5	1·3
8	4·5	4·3	4·5	6·0
9	0	5·3	0	8·0
10	2·3	4·7	2·0	5·5
11	1·3	1·0	1·3	2·2
12	0·8	1·0	0	1·2
13	2·0	5·0	2·0	6·2
Mean	3·02	4·32	2·41	4·62
S.D.	2·60	2·92	2·07	2·85

Table 1*b*. *Mean persistence of the Coriolis sensation following three sets of fixed-interval* (30 *sec.*) 45° *tilt and return motions*

	Counterclockwise rotation			
	7·5 r.p.m.		10 r.p.m.	
Subject	Tilt (S_1)	Return (S_2)	Tilt (S_1)	Return (S_2)
1	6·8	10·7	9·0	11·2
2	4·0	4·0	2·7	3·8
3	6·3	6·5	7·7	9·2
4	7·3	7·8	7·0	7·5
5	5·3	8·7	4·5	9·0
6	3·7	5·0	3·5	6·0
7	7·7	14·2	8·5	13·3
Mean	5·87	8·12	6·13	8·57
S.D.	1·58	3·49	2·53	3·17

The findings from the variable-interval measurements at 7·5 and 10 r.p.m. are shown in Tables 2*a*, *b*; the former shows the results for the OGI subjects, while the latter gives those for the sensation subjects. Since the mean persistence values for both OGI and sensation subjects were not significantly different at 7·5 and 10 r.p.m., the values obtained at these velocities were averaged to give a typical OGI and a

typical sensation response at each S_1–S_2 interval. These averaged values are shown graphically in Fig. 2. Also shown on this graph are the baseline persistence values for S_2 obtained from the fixed-interval measurements. It can be seen that these baseline values correspond very closely to the asymptotic values obtained in the variable-interval measurements.

Fig. 2 shows that the persistence curves for the OGI and sensation subjects are of the same form; the major point of difference is that the averaged OGI values are

Table 2*a*. *Mean duration of OGI evoked by S_2 as a function of interval S_1–S_2 (n = 13)*

	OGI values					
Interval (sec.)	1	2	4	8	16	32
7·5 r.p.m.	0·4	1·8	2·7	3·3	4·1	4·3*
10 r.p.m.	1·0	1·4	2·9	4·0	4·7	5·1*
Grand mean†	0·67	1·58	2·81	3·63	4·38	4·62
S.D.	1·16	1·76	2·55	2·55	2·94	2·95

* n = 8.

† Average of 7·5 and 10 r.p.m. values.

Table 2*b*. *Mean duration of Coriolis sensation evoked by S_2 as a function of interval S_1–S_2 (n = 7)*

	Coriolis sensation values					
Interval (sec.)	1	2	4	8	16	32
7·5 r.p.m.	2·0	2·0	4·0	5·8	8·2	8·3
10 r.p.m.	1·3	2·5	4·6	6·1	7·5	7·9
Grand mean	1·68	2·32	4·32	6·00	7·86	8·14
S.D.	1·42	2·04	2·44	2·62	3·40	3·51

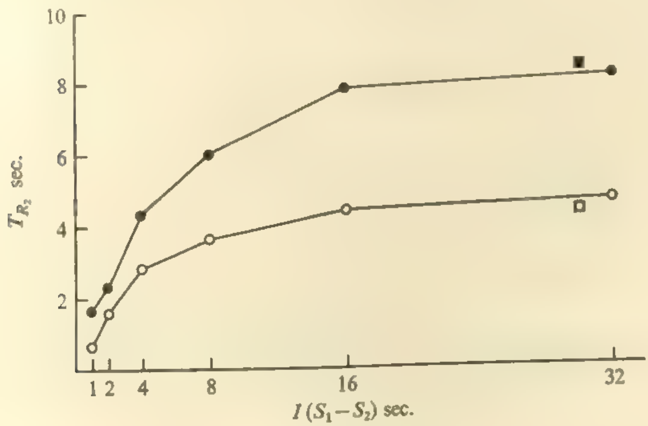


Fig. 2. Mean persistence (T_{R_2}) of the Coriolis sensation and oculogyral illusion (OGI) due to the second stimulus (S_2) as a function of the time interval (I) between the first and second stimuli (S_1–S_2). Also shown are the average sensation and OGI persistence values obtained from the fixed-interval measurements. Both sets of data represent the average of the 7·5 and 10 r.p.m. conditions. ●—●, mean persistence of sensation, variable I; ○—○, mean persistence of OGI, variable I; ■, baseline persistence sensation, fixed I; □, baseline persistence OGI, fixed I.

were [illegible] per cent less than the corresponding sensation values at each test interval. The similarity of these two curves is brought out more clearly by the graphs shown in Fig. 3 in which each mean value for sensation and OGI has been multiplied by a factor chosen to normalize their respective asymptotic values to 10 sec. In the case of the OGI group, this factor was 2·13, whereas for the sensation group, it was 1·25.

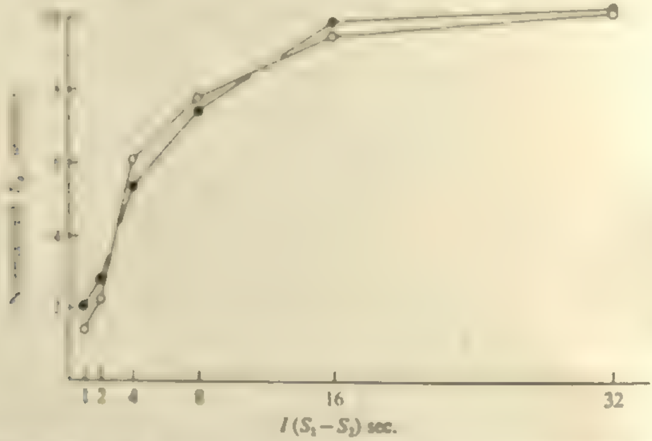


Fig. 3. Normalized Coriolis sensation and oculogyral (OGI) persistence curves. Each mean value for sensation and OGI has been multiplied by a factor chosen to normalize the asymptotic values to 10 sec. ●—●, mean sensation persistence (× 1·25); ○—○, mean OGI persistence (× 2·13).

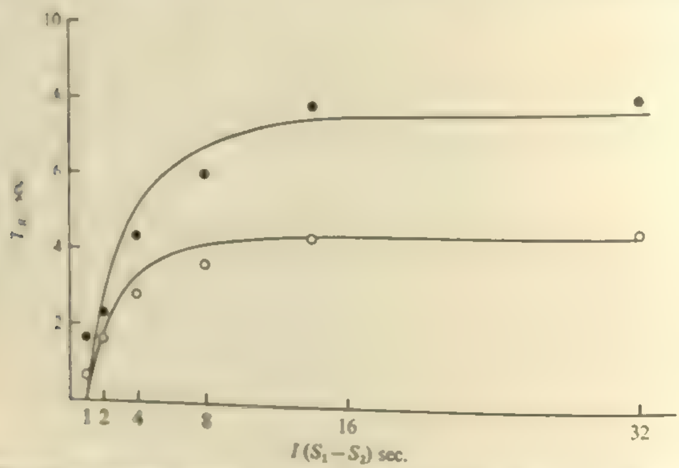


Fig. 4. Comparing observed Coriolis sensation and oculogyral illusion durations with those predicted from the theoretical equation. ●, sensation persistence; ○, OGI persistence; —, theoretically derived curves. (See text.)

The next consideration was how well these observed OGI and sensation persistence values corresponded to the theoretical expectations of the torsion pendulum model. Fig. 4 shows an attempt to fit the observed persistence values for the OGI and sensation by theoretical curves derived from the predictive equation. This was achieved by manipulating the two unknowns, D and k, until an optimum fit was

...ted for each set of obtained values. For the sensation plot, the theoretical curve was fitted with $D = 0{\cdot}192$ and $k = 0{\cdot}212$ where $1/k = 4{\cdot}72$ sec. In the case of the OGI values, the theoretical curve was fitted with $D = 0{\cdot}257$ and $k = 0{\cdot}298$ where $1/k = 3{\cdot}36$ sec. It can be seen that while the theoretical points at the x intercept and the asymptote approximated the observed values, those in the mid range of I intervals tended to overestimate the observed points for both the sensation and the OGI curves.

Discussion

Although incidental to the main purpose of this experiment, it is interesting to note that the subjective responses to the return motion (S_2) were significantly longer than those for the preceding tilt motion (S_1) during the initial fixed interval measurements. During CCW rotation, the tilt motion produces a sensation of pitching backwards or an upward movement of the OGI target, while the return motion has the opposite effects. Guedry & Montague (1961) in their investigation of the psychophysics of the Coriolis reactions noted a similar disparity in the subjective magnitude of certain responses associated with pitch-down and pitch-up sensations elicited by the same type of stimulus. In a more recent study, the present authors (Reason & Graybiel, 1969) also found that magnitude estimates of pitch-down sensations were consistently greater than those for pitch-up sensations produced by Coriolis stimuli of comparable magnitude.

Inasmuch as the theoretical curves (see Fig. 4) derived from the torsion pendulum model approximate fairly closely the obtained OGI and sensation persistence values, the present findings appear to support the *a priori* assumption that the neural events underlying the perception of these subjective phenomena are closely related to the mechanical events occurring within the cupula-endolymph system. However, an entirely peripheral explanation would be inadequate to account for: (1) the finding that the estimated time constants of signal decay (4·72 sec. for sensation and 3·36 sec. for OGI) were considerably shorter than those expected on the basis of theoretical equations governing the mechanics of the canal system (Van Egmond *et al*., 1949); and (2) the finding that the persistence of the Coriolis sensation was proportionately longer than that for the OGI at each time interval.

Theoretical analyses of the Coriolis vestibular reaction (Guedry & Montague, 1961; Lansberg, 1955) have shown that the net effect of the type of stimulus used in this experiment is equivalent to that of a simple impulsive stimulus with the head fixed relative to the axis of rotation. With a 45° head movement, however, the magnitude of the cupula deflexion will be somewhat less than that produced by a simple angular impulse at corresponding angular velocities. Although there is some relative reduction in response magnitude, it is insufficient to account for the greatly reduced time constants obtained in this and similar investigations involving Coriolis stimulation (Guedry & Montague, 1961; Guedry *et al*., 1964). The rate at which the cupula deflexion is annulled by its intrinsic restoring couple should remain constant irrespective of the magnitude of the initial deflexion.

A more probable explanation of this phenomenon is one which postulates some intervention by central mechanisms. Both during and immediately after the 'Coriolis evoking' head motion, the sensory inputs from the canals and otoliths are in

conflict, whereas in the case of the simple impulsive stimulus, the two inputs are normally synergistic. This suggests that the canal signal is suppressed when contradictory graviceptor information is present at the same time. This hypothesis receives additional support from the findings of experiments involving rotation around an Earth-horizontal axis (Benson & Bodin, 1966; Guedry, 1965) and those involving the repositioning of the subject following rotation around the vertical axis (Benson & Bodin, 1965). In both cases, the nystagmus and sensation time constants were considerably shorter than those obtained from simple impulsive stimuli of equivalent magnitude.

The proportionately greater persistence of the Coriolis sensation poses a more difficult problem of explanation since the results from conventional cupulometric studies tend to indicate the contrary finding. Van Dishoeck *et al.* (1954), using averaged cupulograms, showed that the OGI is consistently more persistent than the sensation of rotation over the same range of impulse values.

The values of D and k used to generate the best-fitting theoretical curves (see Fig. 4) suggest that the time constant of the signal mediating the OGI was somewhat shorter and the threshold level higher than that for sensation. Since both phenomena stem initially from the same end-organ activity, it is unlikely that these values reflect the mechanical events within the cupula-endolymph system; more probably, they reflect differences in the transducing properties of the central pathways mediating the two subjective responses.

An additional factor, and one which may also explain the contradictory findings with respect to cupulometry, is that the OGI target used in the present experiment was a relatively large 6-in. cube, whereas in most of the cupulometric studies the target has been a much smaller source of light. It seems reasonable to suppose that a relatively large target having a clearly definable geometric form is more resistant to apparent motion than a smaller, less-defined target. Thus the inherent 'stability' of the present OGI target is likely to manifest itself in a reduced duration of sensation.

Finally, it is worth noting that these findings suggest a technique for minimizing the undesirable consequences of Coriolis stimulation during the early preadapted stages of exposure to a rotating environment. To obtain the greatest advantage from the 'damped swing door' characteristics of the cupula-endolymph system, a necessary head motion should be followed as rapidly as possible by an equal and opposite movement. The present results show that even if this complementary motion is delayed for as long as 4 sec., there is about a 40 per cent loss of sensation persistence with respect to the maximum. For intervals of less than 4 sec., the persistence curve drops precipitately so that at 2 sec. there is approximately a 75 per cent loss, and at 1 sec. the reduction is in the region of 80–100 per cent of the maximum value. It seems reasonable to assume that these persistence losses would also reflect corresponding reductions in the initial magnitude of the subjective phenomena.

This research was conducted under the sponsorship of the Office of Advanced Research and Technology, National Aeronautics and Space Administration, Order No. R-93. We gratefully acknowledge the advice given to us by Dr Alan Benson of the Royal Air Force Institute of Aviation Medicine, Farnborough, England. In addition, we should like to record our indebtedness to Mr Roy Garlock, who assisted in the collection of data.

References

Benson, A. J. & Bodin, M. A. (1965). The effects of the direction of a linear acceleration vector on postrotational vestibular responses in man. (AGARD Conference Proceedings Series no. 2, 9-22.)

Benson, A. J. & Bodin, M. A. (1966). Interaction of linear and angular accelerations on vestibular receptors in man. *Aerospace Med.* **37**, 144.

Graybiel, A., Clark, B. & Zarriello, J. J. (1960). Observations on human subjects living in a 'slow rotation room' for periods of two days. *Archs Neurol.* **3**, 55-73.

Groen, J. J. (1957). Cupulometry. *Laryngoscope* **67**, 894.

Guedry, F. E., Jr. (1965). Orientation of the rotation axis relative to gravity: its influence on nystagmus and the sensation of rotation. *Acta oto-lar.* **60**, 30.

Guedry, F. E., Jr. & Montague, E. K. (1961). Quantitative evaluation of the vestibular Coriolis reaction. *Aerospace Med.* **32**, 487.

Guedry, F. E., Jr., Collins, W. E. & Graybiel, A. (1964). Vestibular habituation during repetitive complex stimulation: a study of transfer effects. *J. appl. Physiol.* **19**, 1005-1015.

Lansberg, M. P. (1955). On the origin of unpleasant sensations elicited by head movements during after-sensations. *Aeromed. Acta* **4**, 67.

Reason, J. T. & Graybiel, A. (1969). Magnitude estimations of Coriolis sensations. (NAMI-1082. NASA Order R-93. Naval Aerospace Medical Institute, Pensacola, Fla.)

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Van Dishoeck, H. A. E., Spoor, A. & Nijhoff, P. (1954). The optogyral illusion and its relation to the nystagmus of the eyes. *Acta oto-lar.* **44**, 597.

Van Egmond, A. A., Groen, J. J. & Jongkees, L. B. W. (1949). The mechanics of the semicircular canal. *J. Physiol., Lond.* **110**, 1-17.

(*Manuscript received* 17 *March* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 175–185
Printed in Great Britain

OCULAR DOMINANCE: A TEST OF TWO HYPOTHESES

By D. M. A. GRONWALL and H. SAMPSON
Department of Psychology, University of Auckland, New Zealand

Fifty subjects were given a battery of eye dominance tests to investigate whether there are five types (Lederer, 1961) or two types (Walls, 1951) of ocular dominance, and to investigate the relation between these tests and handedness. There was no evidence in support of either classification, or of any correlation between the preferred eye and the preferred hand. An alternative hypothesis to those formulated by Lederer and by Walls is presented, which seems to account more adequately for the results.

There have been two theoretical accounts of ocular dominance in which the hypothesis is that eye dominance does *not* represent one single property. One of these is attributable to Lederer (1961), who presents what is perhaps more a classification than a hypothesis. He distinguishes *five* types of ocular dominance 'whose inter-relationships are essentially obscure' (Lederer, 1961, p. 534): Type 1—monocular sighting and aiming. Type 2—motor dominance of one eye in binocular vision. Type 3—orientational dominance, or position of the binoculus. Type 4—sensory dominance. Type 5—dominance of one half of the visual field.

Lederer, in addition, makes two testable propositions about these categories. First, he considers there is no evidence for combining the three motor aspects of dominance (Types 1–3) and, secondly, that monocular sighting could be determined by *either* sensory dominance, *or* motor dominance, *or* orientational dominance.

A more specific hypothesis comes from Walls (1951). In his view there are two entirely independent categories of eye dominance, sensory–perceptual dominance and motor dominance, neither of which is related to handedness or visual acuity. His hypothesis is that the 'essence of motor ocular dominance is the employment of the record of the innervation to the muscles of one eye only for the construction of binocular percepts of visual direction' (Walls, 1951, p. 400). He maintains that the record of the motor dominant eye *alone* determines the perceived direction of objects in relation to the observer, whether he uses binocular or monocular vision. Motor dominance is therefore *directional* dominance.

Walls envisages two subdivisions of the motor category, in which dominance is a *consequence* of the primary directional dominance: (*a*) *secondary directional dominance* which refers to the eye which diverges less, or not at all, when covered after a bifixation target has been brought and held 3 in. from the face, and (*b*) '*master eye*' *dominance* which is the eye before which one holds a card to read.

In the case of secondary directional dominance, it follows from his hypothesis that, if an innervation record is kept of the dominant eye only, then the non-dominant eye can move under cover without any change in apparent target position. But even under cover the dominant eye must maintain steady fixation, or the target will appear to be changing in direction. In the case of 'master eye' dominance, the dominant eye behaves as if it were lazy. If every movement of the dominant eye is important in spatial localization, then it seems logical that this eye will not make

unnecessary movements, and all the vergences required for focusing will be left to the non-dominant eye.

Although several workers have correlated different eye dominance tests (e.g. [illegible], 1947; Crider, 1935; Crovitz, 1961), there does not appear to be any investigation in which a sufficient range of tests has been administered to evaluate Lederer's classification or Walls' hypothesis. Yet they both have serious implications for the use of eye dominance as a measure of laterality. It was the purpose of this study, therefore, to attempt this evaluation.

To summarize, the purpose of the present study was to administer a representative range of eye dominance tests in order to answer the following questions. (1) Do the tests measure a unitary property, 'eye dominance'? (2) Are there at least two types of eye dominance (Lederer, 1961)? (3) Are there two different kinds of eye dominance (Walls, 1951)?

METHOD

Subjects

Fifty students between 17 and 22 years old, from the first-year psychology course at Auckland University, were given the complete test battery. All had *uncorrected* binocular vision of 20/20 or better. One subject was discarded, and another tested in his place, when he proved unable to see the [illegible] phi movement.

The subjects were not told the purpose of the experiment; merely that they would be given a number of tests involving vision. All tests were administered individually.

Apparatus and procedure

For administration, the tests* were divided into four groups.

Group I. Simple tests of eye dominance. (i) The Card Test (Crider, 1944). (ii) The Ring Test (Crider, 1944). (iii) Miles' A-B-C Test (Miles, 1929). (iv) The Box Test (Crider, 1944). (v) The [illegible] Rosenbach Test (cited by Lederer, 1961). (vi) Monocular visual acuity (Snellen Chart). (vii) Hold Card to Read (Walls, 1951). (viii) Eye closing (Crider, 1944). (ix) Mills' Test (Crider, 1944). (x) The Cover Test. While the subject bifixates a target held 3 in. from the face, each eye in turn is covered with a small magnifying glass. The eye which diverges less or not at all under cover is recorded. (xi) Handedness questionnaire (Humphrey, 1951, with three additional items by [illegible] *et al.*, 1960). (xii) Facing a surface square (Walls, 1951).

Group II. Stereoscopic rivalry. A table-stand stereoscope (Keystone Correct-Eye-Scope) was used with specially designed cards for the purpose.

Group III. Phi movement. The apparatus was modelled on that described by Carter (1953) for his three-target situation.

Group IV. Psychophysical methods. (i) Alignment. The apparatus was a modification of that reported by [illegible] (1942), and suggested by Lederer (1961) as a suitable measure of directional dominance. (ii) Chromatic test. Lederer (1961) cites Pascal's chromatic brilliance test which was used here.

Order of administration of the tests

Each testing session was divided into two half-hour blocks, one block beginning with the phi test (group III), and the other the group IV psychophysical tests. These were both followed by six of the group I tests, alternating with six stereoscopic rivalry cards, all given in a different random order for each subject. Odd-numbered subjects began the session with the group III, and even-numbered subjects with the group IV block.

* See appendix.

Results and discussion

1. *Do the tests measure a unitary property, 'eye dominance'?*

Duke-Elder (1949), using sighting tests, estimated that 65 per cent of the population is right-eyed and 35 per cent left-eyed. Similar percentages were reported in a survey of [illegible] school pupils by Spong (1962), who also used a sighting test as an index of eye dominance and found among each age group of children in his survey (7–9, 10–12 and 13–16 years) that there were approximately 65 per cent right-eyed and

Table 1. *Total subjects classified by each test as right or left dominant* ($n = 50$) *compared with the population sighting eye dominance figures of* 65 *per cent right and* 35 *per cent left eyed* (*Duke-Elder*, 1949, *Spong*, 1962)

Test	Right	Left	χ^2 Frequency†	χ^2 Proportion‡
Card	30	18	< 1	< 1
Ring	29	17	< 1	< 1
Box	27	16	1·06	< 1
Miles	27	20	1·29	< 1
Porta Rosenbach	34	14	< 1	< 1
Acuity	12	5	21·86**	< 1
Eye closing	29	16	< 1	< 1
Hold card	35	13	1·35	< 1
Mills'	11	18	14·23**	5·74*
Cover test	15	18	9·43**	3·18
Handedness	43	4	—	—
Face squarely	20	19	4·94*	1·68
Rivalry: colour§	4	4	35·40**	< 1
Rivalry: form§	2	3	40·63**	< 1
Rivalry: word§	14	6	18·09**	< 1
Alignment	7	12	21·74**	4·18*
Chromatic	11	11	16·63**	1·15

* $P < 0·05$ ** $P < 0·01$.

† Comparison of the obtained frequency with the expected frequencies of 32·5 right-eyed and 17·5 left-eyed subjects.

‡ Comparison of the proportion of right to left cases with the expected proportion of 65:35.

§ Subjects showing holocular dominance only.

35 per cent left-eyed. If all the tests in the present study measure a common property, the frequencies of right- and left-eyed subjects might be expected to approach these frequencies. χ^2 comparisons of the obtained *frequencies* of eye dominance on each test with these expected frequencies are given in column 4 of Table 1. Table 1 gives all instances where subjects were classified eye dominant and not ambiocular on each test. Nine of the 16 tests differed significantly from the 35·5 right-eyed and 17·5 left-eyed cases that would be expected in this sample. This suggests that more than one property is being measured by the tests.

However, it must be remembered that both figures for the percentage of eye dominance in the population result from the administration of *sighting* tests. It may be, therefore, that there is only *one* kind of eye dominance, and that all the tests are not equally effective in detecting it. Some tests may only detect a proportion of subjects classified by sighting tests, *but the trait they measure may be the same*. If this is so, and all the cases of eye dominance that *are* detected come from a common population, then the proportion of right- to left-eyed subjects should be 65:35. χ^2

Table 2. *Correlations between the dominance tests administered in the test battery to 50 subjects*

(τ_b; Kendall, 1955)*

	2 Ring	3 Box	4 Porta–Rosenbach	5 Miles	6 Eye closing	7 Hold card	8 Mills' test	9 Cover test	10 Acuity	11 Rivalry: form†	12 Rivalry: colour†	13 Rivalry: word†	14 Alignment	15 Phi: half field	16 Handedness	17 Chromatic	18 Face
1 Card	0·83	0·70	0·66	0·60	0·47	0·41	0·28	0·17	0·33	0·34	0·20	0·19	0·14	−0·05	0·06	0·06	0·01
2 Ring	—	0·76	0·69	0·56	0·40	0·37	0·38	0·29	0·29	0·33	0·24	0·23	0·14	−0·04	0·12	0·08	0·02
3 Box	—	—	0·69	0·51	0·38	0·30	0·26	0·15	0·27	0·35	0·36	0·25	0·23	0·09	0·01	0·16	0·05
4 Porta–Rosenbach	—	—	—	0·44	0·44	0·34	0·34	0·18	0·15	0·37	0·21	0·36	0·31	0·18	0·01	−0·06	−0·03
5 Miles	—	—	—	—	0·25	0·47	0·37	0·27	0·13	0·35	0·36	0·35	0·22	0·01	0·12	0·03	−0·07
6 Eye closing	—	—	—	—	—	0·29	0·19	0·00	0·09	0·08	0·07	0·07	0·21	0·14	−0·07	−0·11	0·15
7 Hold card	—	—	—	—	—	—	0·11	0·02	0·35	0·38	0·22	0·19	0·42	0·01	0·01	−0·09	−0·09
8 Mills' test	—	—	—	—	—	—	—	0·41	−0·19	0·12	0·06	0·17	0·15	−0·04	0·07	0·01	−0·02
9 Cover test	—	—	—	—	—	—	—	—	0·04	0·11	0·11	0·13	0·12	0·04	0·10	0·16	−0·09
10 Acuity	—	—	—	—	—	—	—	—	—	0·34	0·16	0·23	0·19	−0·05	0·03	−0·14	−0·23
11 Rivalry: form†	—	—	—	—	—	—	—	—	—	—	0·47	0·50	0·49	−0·11	0·17	−0·09	0·002
12 Rivalry: colour†	—	—	—	—	—	—	—	—	—	—	—	0·27	0·24	0·08	0·00	0·00	0·00
13 Rivalry: word†	—	—	—	—	—	—	—	—	—	—	—	—	0·39	−0·01	0·18	−0·09	−0·03
14 Alignment	—	—	—	—	—	—	—	—	—	—	—	—	—	−0·01	−0·07	−0·19	−0·01
15 Phi: half-field	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0·03	−0·12	−0·11
16 Handedness	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0·09	0·12
17 Chromatic	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	−0·12

* Correlations over 0·19 and 0·25 are significant at the 0·05 and 0·01 levels respectively.

† Holocular dominance only.

comparisons of the number of right- and left-classified subjects on each test with the expected proportion of 65 : 35 are given in column 5 of Table 1. Of the 16 tests, only Mills' and the alignment test detect significantly different proportions of right-eyed and left-eyed cases. Although this does not *prove* that eye dominance is a unitary trait, the possibility that all the tests *are* measuring the same thing, some more efficiently than others, remains.

Table 3. *Mean and maximum τ_b correlations* within and between groups selected by a cluster analysis from the matrix of Table 2*

Test no †		1–5	6	7	8, 9	10–13	14	15–18
1–5	Mean	0·65	0·39	0·38	0·27	0·28	0·21	0·065
	Maximum	0·83	0·47	0·47	0·38	0·37	0·31	0·18
6	Mean	—	—	0·29	0·09	0·08	0·21	0·02
	Maximum	—	—	0·29	0·19	0·09	0·21	0·15
7	Mean	—	—	—	0·07	0·29	0·42	−0·03
	Maximum	—	—	—	0·11	0·38	0·42	0·01
8, 9	Mean	—	—	—	0·41	0·06	0·14	0·005
	Maximum	—	—	—	0·41	0·17	0·15	0·16
10–13	Mean	—	—	—	—	0·33	0·33	−0·02
	Maximum	—	—	—	—	0·47	0·49	0·18
14	Mean	—	—	—	—	—	—	−0·07
	Maximum	—	—	—	—	—	—	−0·01

* Mean τ_b value for tests within cluster from Table 2 given in top row, and maximum correlation in each group in bottom row.

† Numbers refer to tests in Table 2.

Table 2 gives the matrix of correlations between the tests in the battery. Every test classified subjects as right-eyed, ambiocular or left-eyed. The τ_b correlation statistic (Kendall, 1955) was used. Each correlation in Table 2 uses data from all 50 subjects. (With $n = 50$, τ_b values above 0·19 are significant at the 5 per cent level, and above 0·25 at the 1 per cent level of confidence.)

A cluster analysis (Adcock, 1952) was performed on the correlation matrix. It can be seen from Table 2 that tests 1–5, which form the nucleus of the main cluster, are all significantly correlated ($P < 0{\cdot}01$). The criterion for exclusion of tests from a group was the occurrence of insignificant or negative correlations between them. Tests 15–18 are clearly excluded from every group. The same data are presented in a simplified version in Table 3, which shows only mean and maximum τ_b values within and between groups.

Obviously there would be clear evidence for only one type of eye dominance if all the tests were interrelated. This does not occur, as the following statements of the relationships given in Tables 2 and 3 show: (*a*) Tests 1–5 correlate highly with each other and moderately with all tests except 15–18. We will refer to these as the nuclear tests. (*b*) The two eye-muscle tests (8, 9), themselves significantly correlated, correlate moderately with the nuclear tests, but not with tests 10–14. (*c*) Tests 10–13, the sensory tests, correlate significantly with each other and with the nuclear tests. (*d*) The eye-closing test correlates with the nuclear tests, but not with sensory or muscle tests. (*e*) The hold-card test correlates with the nuclear tests and the sensory tests, but not with the muscle tests. (*f*) The alignment test correlates with only three

of the nuclear tests with the sensory, eye-closing and hold-card tests, but not with the muscle tests. (c) Tests 15–18 show no correlation with each other, or with any of the above tests.

This does not give unequivocal evidence. Statement (*a*) above suggests that all tests except 15–18, are measuring a common property. The other statements suggest that some tests are measuring quite different things.

In summary, the evidence considered here does not give an unequivocal answer to the question: is there only one kind of eye dominance?

2. *Are there at least five types of eye dominance?*

Table 4 shows the tests grouped into five types according to Lederer's (1961) classification. This table is also a simplification of the complete matrix in Table 2, and only the mean τ_b correlations within and between groups of tests, together with the maximum value shown in each case, are given. The hold-card results are not included, as Lederer (1961) makes no reference to this test.

Table 4. *Mean and maximum τ_b correlations* within and between groups of tests when grouped according to the classification suggested by Lederer* (1961)

	Sighting Type I	Motor Type II	Orientational Type III	Sensory Type IV	Cerebral Type V	Dubious†
Test no.‡ ...	**1–5**	**8, 9**	**14, 18**	**10–13, 17**	**15, 16**	**6**
Type I						
Mean	0·65	0·27	0·10	0·22	0·05	0·39
Maximum	0·83	0·38	0·31	0·37	0·18	0·47
Type II						
Mean	—	0·41	0·04	0·07	0·04	0·10
Maximum	—	0·41	0·15	0·17	0·10	0·19
Type III						
Mean	—	—	−0·01	0·13	−0·02	0·18
Maximum	—	—	−0·01	0·49	0·12	0·21
Type IV						
Mean	—	—	—	0·18	0·01	0·08
Maximum	—	—	—	0·50	0·08	0·09
Type V						
Mean	—	—	—	—	0·03	0·04
Maximum	—	—	—	—	0·03	0·14

* Correlations from Table 2 were averaged for all tests within each group; for comparison the maximum value of the correlations included in each mean score is shown.
† Test specified by Lederer as of doubtful value in determining eye dominance.
‡ Numbers refer to tests in Table 2.

It is clear from Table 4 that correlations obtained in this study do not justify five different types of eye dominance. Although Lederer does not specify that the five types are unrelated, five separate groups can surely only be justified if the within-group correlations exceed those between groups. This condition is not met. It should be noted that every test used in the classification was specified by Lederer (1961) as an example of the type of eye dominance in which it was classified.

Lederer's proposition that there are three kinds of motor dominance also could not be verified. The maximum correlations between a Type I (sighting) and either a

Type II (motor) or a Type III (orientational) test are all significant at the 1 per cent level, and are very little below the correlation *between* the motor tests within Type II. Nor is there evidence for the prediction that the basis of sighting dominance will be found in either sensory, motor *or* orientational dominance. The correlations of these three groups with the sighting tests are not sufficiently different to justify the proposition that one, but not all, of them can be the basis of sighting dominance.

In summary, the classification and propositions concerning ocular dominance which Lederer (1961) has suggested have not been verified by the present study.

3. *Are there two different kinds of eye dominance?*

Table 5 shows the tests grouped as predicted by Walls' (1951) theory. The data are taken from the correlation matrix of Table 4, simplified by showing only mean and maximum values within and between groups of tests. Walls (1951) does not mention the chromatic test, which theoretically could belong with either the acuity or the rivalry data. It has therefore not been included in Table 5.

Table 5. *Mean and maximum τ_b correlations* within and between tests when grouped according to the classification suggested by Walls* (1951)

	Sensory Type I	Directional Type II	Secondary II*a*	Master eye II*b*	Dubious†
Test no.‡ ...	11–13	1–5, 14	8, 9	6, 7, 18	10, 16
Type I					
Mean	0·41	0·31	0·12	0·12	0·18
Maximum	0·50	0·49	0·17	0·38	0·34
Type II					
Mean	—	0·50	0·25	0·25	0·13
Maximum	—	0·83	0·38	0·47	0·33
Type II*a*					
Mean	—	—	0·41	0·04	0·005
Maximum	—	—	0·41	0·19	0·10
Type II*b*					
Mean	—	—	—	0·10	0·08
Maximum	—	—	—	0·29	0·35
Dubious					
Mean	—	—	—	—	0·03
Maximum	—	—	—	—	0·03

* Correlations from Table 2 were averaged for all tests within each group; maximum value of the correlation included in the mean score is given for comparison.

† Tests specified by Walls as of little value in determining eye dominance.

‡ Numbers refer to tests in Table 2.

It can be seen from the correlations given in the table that they do not support the two different types of ocular dominance hypothesized by Walls. The mean correlation (0·31) between Type I and Type II tests is not lower than that between Type II and its subgroups II*a* and II*b* (both 0·25). Also, the maximum correlation between a Type I and a Type II test is no lower than the maximum correlations of either II*a* or II*b* tests with those of Type II.

The 'dubious' category *does* contain cases of significant correlation with the other dominance measures. Indeed, the maximum correlation (0·35) of the 'dubious' group with the 'master eye' group is higher than the maximum correlation (0·29) between

the two 'master eye' tests. Finally, Walls asserts that secondary directional dominance and 'master eye' dominances are *both* consequences of a primary directional dominance. Yet these two subgroups are themselves insignificantly correlated. Once again, all tests used here were either explicitly or implicitly suggested by Walls (1951) as examples of the dominance types he postulates.

In summary, the data of the present study fail to confirm the hypothesis proposed by Walls (1951) concerning the nature of ocular dominance.

CONCLUSIONS

It is clear that neither Lederer's nor Walls' hypotheses account for these data satisfactorily. However, the present experiment does lead to a number of conclusions that suggest a possible answer to the question of the nature of eye dominance.

1. *No division of eye dominance tests into independent classes can be supported.* Five of the tests (the nuclear group) correlated highly with each other and moderately with all other tests in the battery (the peripheral group). Further, as was argued earlier, analysis of the proportion of right- to left-eyedness on the peripheral tasks is consistent with the proposal that these tests detected cases from the same population of eye preference as the nuclear group, but less efficiently, in that they classified fewer cases.

2. *Eye preference refers to the consistent use of one eye for sighting and aiming.* In each of the five nuclear tests information from only one eye must be used. Although instructions require the subject to keep both eyes open, efficient performance in these sighting and aiming tasks depends on the subject ignoring binocular vision. It is evident from the present study that equating eye preference with the use of one eye for sighting and aiming is appropriate, because the nuclear group clearly forms the core of the eye preference tests. 'Eye preference' is suggested as a more suitable term than 'eye dominance', which implies some intrinsic superiority of one eye. It appears to result from the establishment of motor control over the binocular visual system to convert it to a monocular one to meet the demands of a visual task.

3. *The peripheral group are primarily tests of visual system asymmetries.* The significant correlations between the nuclear and the peripheral tests suggest that the latter are also measuring eye preference. However, they are not sighting tests, and from present results cannot be considered to measure different kinds of eye preference. The most parsimonious explanation is that they are primarily tests of visual system asymmetries that on occasion may influence which eye becomes the preferred one.

4. *Consistent eye preference is measured reliably only when the use of one eye rather than the other does not seem to the subject to be a relevant feature of his performance.* The alignment and the Porta–Rosenbach tests apparently set the subject similar tasks: lining up a close with a far object by ignoring one of the diplopic images of the former. There was, however, only a moderate correlation between the two tests, and a crucial difference appears to be that, while only two trials of the Porta–Rosenbach were given, in the alignment test the subject had a 5 min. practice session followed by 15 recorded trials. The discrepancy between the results of the two tests has implications for the measurement of eye preference, and also suggests a reason for disagreement among investigations (e.g. Buxton & Crosland, 1937; Crider, 1944). All

rapidly administered tests can be given without the subject's knowledge of the real purpose of the exercise. However, as the number of repetitions of one test increases, so does the probability that the subject will become aware that eye preference is being measured. That giving different numbers of trials on a test may influence results is also implied in Clark's (1952) suggestion that the data of any two investigations are not comparable unless subjects have been given an equal number of sighting opportunities.

References

Adcock, C. J. (1952). *Factor Analysis for Non-Mathematicians*. Wellington: Victoria University College.

Buxton, C. E. & Crosland, H. R. (1937). The concept 'eye-preference'. *Am. J. Psychol.* **49**, 458–461.

Carter, D. B. (1953). A further demonstration of phi movement cerebral dominance. *J. Psychol.* **36**, 299–309.

Clark, M. M. (1952). Measurement and interpretation of eye dominance. *Nature, Lond.* **170**, 192–194.

Crider, B. (1935). The relation of eye muscle balance to the sighting eye. *J. exp. Psychol.* **18**, 152–154.

Crider, B. (1941). Eye closure facility and eye dominance. *J. genet. Psychol.* **58**, 425–426.

Crider, B. (1944). A battery of tests for the dominant eye. *J. gen. Psychol.* **31**, 179–190.

Crovitz, H. (1961). Differential acuity of the two eyes and the problem of ocular dominance. *Science, N.Y.* **134**, 614.

Diehl, H. T. (1942). An eye dominance gauge and some of its uses. *J. gen. Psychol.* **28**, 181–184.

Duke-Elder, W. S. (1949). *Textbook of Ophthalmology*, vol. 4. London: Kimpton.

Gillies, S. M., MacSweeney, D. A. & Zangwill, O. L. (1960). A note on some unusual handedness patterns. *Q. J. exp. Psychol.* **12**, 113–116.

Humphrey, M. E. (1951). Consistency of hand usage: a preliminary enquiry. *Br. J. educ. Psychol.* **21**, 214–225.

Kendall, M. G. (1955). *Rank Correlation Methods*, 2nd ed. London: Griffin.

Lederer, J. (1961). Ocular dominance. *Aust. J. Optom.* **44**, 531–574.

Miles, W. (1929). Ocular dominance demonstrated by unconscious sighting. *J. exp. Psychol.* **12**, 117–127.

Spong, P. (1962). *Recognition and Recall of Retarded Readers: a Developmental Study*. Winifred Gimblett report. University of Auckland, N.Z.

Walls, G. L. (1951). A theory of ocular dominance. *Archs Ophthal.* **45**, 387–412.

(*Manuscript received* 25 *February* 1970)

APPENDIX

Eye dominance tests used in the study

Group I. Rapid tests of eye dominance

(i) *The Card Test* (Crider, 1944). The subject holds a card in front of his face with both hands and with both eyes open looks through a small hole at a coloured spot on a card held by the experimenter. This is repeated with three different cards, and dominance recorded if the same eye is used with each card.

(ii) *The Ring Test* (Crider, 1944). The subject fixates a point, then lifts up a 60 mm diameter ring until he can see that point through it. This is repeated with the subject holding the ring in his non-preferred hand, and dominance recorded if the same eye is enclosed by the ring on both occasions.

(iii) *Miles' A-B-C Test* (Miles, 1929). The subject covers his face with a truncated white cardboard cone, which has to be held open with two hands if he is to see through the aperture. He is asked to keep both eyes open and state which of two coloured patches on a card held by the experimenter is darker. Five different cards and cones are used, and dominance recorded if the same eye is used on at least four trials.

(iv) *The Box Test* (Crider, 1944). The subject lines up two pipecleaners inserted in the open ends of a cardboard tube held 6–8 in. in front of his face. Two trials are given, and dominance recorded if the same eye is used for aligning on both occasions.

(v) *The Porta–Rosenbach Test* (cited by Lederer, 1961). The subject holds a pencil at arm's length with two hands, and lines it up with a small mark on a wall approximately 10 ft. from his eyes. He is asked to close each eye alternately and report with which eye the alignment is effected. This is repeated once, and dominance recorded if the same eye is used each time.

(vi) *Monocular visual acuity.* Acuity is checked for each eye separately on a Snellen chart. All subjects having 20/20 vision in both eyes are classified ambiocular.

(vii) *Hold Card to Read.* Four $3\frac{1}{2} \times 3$ in. white cards, each with a different word in fine print, are placed face down on the table before the subject and he is asked to pick each card up and read the words as quickly as possible. A stop-watch is used to convince the subject that speed, rather than the eye being used, is the dependent variable, and dominance recorded if all four cards are held before the same eye.

(viii) *Eye closing.* The subject is asked to close each eye in turn and report which eye is more difficult to close. This is repeated once, and dominance recorded if the subject's report and the experimenter's observations agree on both trials.

(ix) *Mills' Test.* The subject maintains fixation on a spot, which is slowly moved towards his eyes until the near point of convergence is reached and passed.

(x) *The Cover Test.* While the subject bifixates a target held 3 in. from his face, each eye in turn is covered with a small magnifying glass. The eye which diverges less or not at all under cover is recorded.

(xi) *Handedness Questionnaire.* This comprises 20 items suggested by Humphrey (1951), together with three additional items designed by Gillies *et al.* (1960). Subjects with a dextrality index of 0·70 or more are classified right-handed, and of 0·30 or less, left-handed.

(xii) *Facing a surface squarely.* The subject is asked to hold up an 8×7 in. white card about 12 in. from his face so that he is facing it squarely and his line of sight is perpendicular to the plane of the card. This is repeated once, and dominance recorded if the same eye is more nearly perpendicular to the card on both occasions.

Group II. Stereoscopic rivalry

A table-stand stereoscope (Keystone Correct-Eye-Scope) is used with 12 viewing cards, four each colour, form and words. Each card has as fixation point two black dots, 2·5 in. apart, which are centred on the whole-eye stimuli and 1 mm to the left or right of the hemiretinal stimuli. Subjects are instructed to maintain fixation on the dot and to report rivalry changes as they occur over a 60 sec. period with each card, the experimenter activating an electronic timer to record the duration of each type of

response. For each card, eye or hemiretinal dominance is recorded if reports to that area exceed those to the opposing target by 20 per cent or more.

Group III. Phi movement

The apparatus is modelled on that described by Carter (1953) for his three-target situation. An electronic timer is activated as soon as movement in one direction is reported, and the duration of each type of movement recorded for a 2 min. period. Dominance is scored if there is a 20 per cent or greater excess of movement in one direction.

Group IV. Psychophysical methods

(i) *Alignment test.* The apparatus is a modification of that reported by Diehl (1942) and suggested by Lederer (1961) as a suitable measure of directional dominance. The subject looks through a viewing screen, with head immobile, at two 3 in. high white pins. The nearer is fixed in position 6 ft. from the subject and immediately in line with the point midway between the two viewing holes. The second pin runs in a groove parallel to the screen and 7 ft. distant from it, and its position can be adjusted by pulling a continuous cord attached to the base of the pin at both sides. The subject is instructed to line up the pins, and given two practice trials before testing begins. There are three test conditions, given in a different random order for each subject: binocular, with both viewing holes clear, and two monocular conditions with the right and left eye hole in turn covered with a card screen. Five settings are made under each condition, with the variable pin moved to different right and left positions at the beginning of alternate trials. Directional dominance is scored for one eye if the difference between the mean settings for that eye and the binocular condition is less than 0·25 of the difference between the mean settings under the two monocular conditions.

(ii) *Chromatic test.* A coloured filter, red or blue theatrical gelatine mounted in a cardboard transparency frame, is inserted over one eye hole of the viewing screen. The subject is asked to judge whether each of a series of comparison filters, all made of the *same* red or blue material placed over the other eye, is the same as, darker than or lighter than the standard. Five red and five blue comparisons are made with the standard over each eye, with a random order of these four conditions. An eye is scored as dominant if there is more than a 25 per cent excess of darker judgements given to filters over that eye.

Printed in Great Britain

STRATEGIES IN ROTARY PURSUIT TRACKING

By C. D. FRITH

Department of Psychology, Institute of Psychiatry
University of London

Ten measures were derived describing different aspects of the pursuit rotor performance of 30 subjects. These [illegible] variables [illegible] the shapes of the [illegible] and [illegible] lengths and also variables indicating the amount of [illegible] present in performance. The relationship between the various measures suggested that performance [illegible] in terms of two factors. The first represented the level of attainment (total time on target). The second was independent of the first and represented different strategies of performance. The measures defining these strategies suggested that at one extreme people were concerned only with velocity matching, while at the other extreme they were concerned only with position matching. There was a strong relationship between strategy and personality, extraverts adopting velocity matching and introverts adopting position matching. There was no evidence that these differences in response style were due to the greater production of rest pauses by the extraverts.

Most of the changes in performance which occur during massed practice on the pursuit rotor have been explained, at least in part, in terms of Hull's (1943) concept of reactive inhibition (I_R). It has been suggested (Kimble, 1949) that during massed practice the build-up of reactive inhibition eventually produces gaps in performance (rest pauses) resulting in an overall lowering of performance. During a programmed rest the I_R will dissipate, resulting in better performance immediately after the rest (reminiscence). Eysenck (1956) has suggested that extraverts, because of their greater proneness to the build-up of inhibition, should show more rest pauses, and therefore more reminiscence than introverts. (More recently (Eysenck, 1965) he has explained the differences in reminiscence between introverts and extraverts in terms of consolidation and conditionability.) However, no direct evidence for the existence of these rest pauses has yet been produced. Although much study has been made of performance decrement, i.e. post-rest down-swing, in relation to such variables as length of pre-rest practice, length of rest, drive, etc., few detailed investigations of the actual changes in performance involved have been carried out. This is because few studies have ever considered any measure of pursuit rotor performance other than total time on target. This is a measure of attainment which does not allow distinction between different styles, or strategies, of performance which result in the same level of attainment. Differences in strategy might be the crucial component of individual differences in performance. This has already been found to be the case with another very simple motor skill. Spielman (1963) demonstrated that in a tapping task extraverts produced more rest pauses than introverts, but that both types of people produced the same number of taps. In this case simple attainment level could not reveal an important dimension of individual difference, nor could it reveal the existence of rest pauses. These could only be demonstrated in the distribution of intertap intervals. For people producing rest pauses this distribution was bimodal, having a small secondary distribution of abnormally long intertap intervals. To demonstrate the existence of rest pauses in rotary pursuit tracking it would similarly be necessary

to study the distribution of miss lengths (i.e. lengths of times off target). The experiment to be reported here involved precisely this.

However, a previous experiment (Frith, 1968) involving rather simpler measures had already revealed interesting features of pursuit rotor performance which could not have been found from the study of total time on target alone. In this experiment the number of times contact was made with the target (hits) was recorded in addition to the total time on target. From these two measures the average time of each unbroken contact (average hit length) and the average miss length could be estimated. Thus two independent measures of performance were available. The total time on target indicated the attainment level. The number of hits could indicate different strategies of performance for the same total time on target, i.e. either many short hits and misses or few long hits and misses. It was found that while extraverted and introverted people did not differ in attainment level, extraverted people tended to adopt the strategy involving long hits and misses. This finding would be consistent with the hypothesis mentioned earlier that extraverts produce more rest pauses than introverts, if long average miss lengths can be taken to indicate the presence of rest pauses. No definite conclusions can be reached until the distributions of hit and miss lengths have been studied. Rest pauses must be defined as abnormally long misses occurring during performance. Such abnormality can only be defined in terms of the total distribution of misses.

METHOD

A variable pattern polar tracker (Model PRI 15, Shaw Laboratories, N.Y.) was used. In this piece of apparatus the target is provided by a radial strip of light set in a revolving turntable. Above this is placed a sheet of glass, the undersurface of which is covered with light-proof paper. Tracks of any shape can be made by cutting away the appropriate parts of this paper. The target is then seen as a patch of light moving around the track. The stylus was a rigid L-shaped rod with a photo-electric cell at its tip which closed a relay whenever it was over the target patch of light. The maximum speed of revolution of the target was 37·5 r.p.m. The track consisted of a centrally placed equilateral triangle with sides 17·1 cm long and 1·2 cm wide.

Since all points on this track are not equidistant from the centre of revolution the target changes speed, moving faster along those parts of the tracks (the corners) which are farthest from the centre.

The photo-electric pursuit rotor was linked with an oscillator and tape-recorder so that whenever the person was on target a tone was recorded on the tape. The tape was then played back into a LINC 8 computer through a triggering device which caused the tones recorded on the tape to operate a relay. The main program used to analyse these data was called ROTMAP, which processed a 5 min. session of work on the pursuit rotor which it divided into ten 30 sec. periods. For each period the program provided histograms of hit and miss lengths from 50 to 5000 msec. in 50 msec. steps and also the total time on target and the total number of hits.

There were 30 subjects, male volunteers aged between 20 and 35. They worked at the pursuit rotor for three 5 min. sessions separated by 10 min. rest. The first 5 min. session was treated as practice and not recorded, since it was thought that any strategies of performance would take some time to develop and stabilize. All subjects were given the EPI (Eysenck & Eysenck, 1964). The distribution of scores was as follows: mean Extraversion score = 10·2; S.D. = 5·4; range = 2–23; mean Neuroticism score = 10·3; S.D. = 5·9; range = 2–22.

Instructions were kept to a minimum, all the subjects being told to 'keep the end of the stylus on top of the moving light'. A demonstration performance was then given by the experimenter. When the subject was on target a red light on the front of the pursuit rotor apparatus lit up and this was pointed out to all subjects, but they could not hear the tone being recorded on tape.

Results

Description of performance

The following analysis is based on data from the whole of the second 5 min. session of work. The third 5 min. session gave essentially the same results

The most obvious feature of the hit and miss distribution was the large positive skew. The position of the mode was remarkably constant for all subjects, being at 200 msec. for the hit distribution and at 100 msec. for the miss distribution. If only those parts of the curves to the right of the mode were considered they bore a strong resemblance to the Poisson distribution, this being particularly so with the distribution of miss lengths. The Poisson distribution is a particularly attractive curve with which to describe these data since it is completely defined by one parameter. Also,

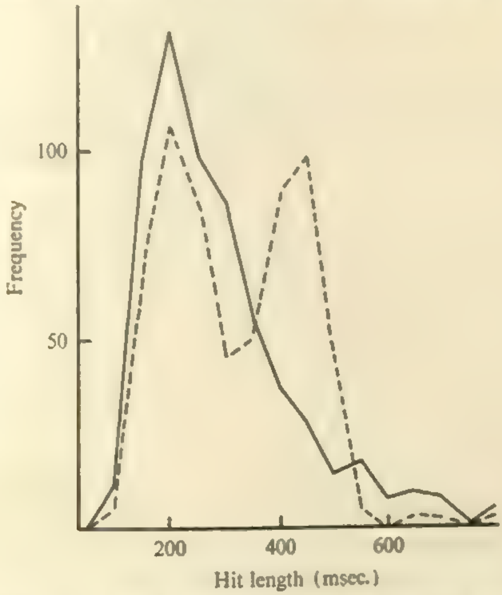


Fig. 1. Markedly different hit-length distributions for two subjects with the same level of performance in terms of total time on target. – – –, Subject H; ——, subject A.

the model itself has an intuitive appeal in the case of pursuit rotor performance, since the defining parameter would be the probability of coming off target in the case of the hit length distribution and the probability of coming on target in the case of the miss length distribution. The probabilities of coming on target and of coming off target should clearly be important components of pursuit rotor performance. They should be closely related to attainment, high attainment being associated with high probability of coming on target and low probability of coming off target. These two probabilities were estimated from the hit distributions. The probability of coming on target at 200 msec. after coming off it was given by the number of times this event occurred (i.e. number of miss lengths of 200 msec.) divided by the number of times the subject came back on target after this time interval or longer (i.e. number of miss lengths $\geqslant$ 200 msec). The Poisson model requires that this probability should be the

same for all miss lengths and therefore only one estimate is needed for each subject. The most accurate estimation will be one involving the largest number of events. Probabilities were therefore calculated from the five classes of hit and miss lengths from 150 to 350 msec. and averaged.

The parts of the distributions to the left of the mode did not fit the Poisson model there being too few short hit and miss lengths. Therefore a second parameter of the distributions was calculated for each subject which was the probability of coming back off or on target within 150 msec. or less.

There was another striking deviation from the Poisson distribution in that for many subjects the hit distribution was markedly bimodal. Fig. 1 shows the hit distribution for two subjects of the same performance level, one showing this bimodality and the other not. The mode of this secondary peak was also relatively constant across subjects, occurring at between 400 and 450 msec. Therefore a measure indicating the amount of bimodality in the hit distribution was calculated for each subject by subtracting the number of hits of length 300 and 350 msec. from the number of hits of length 400 and 450 msec., a high degree of bimodality giving a positive score and a lack of bimodality a negative score. It was considered that these five measures (Table 1) would sufficiently describe the hit and miss length distributions for each subject.

Five other measures were also derived to describe other aspects of performance. One was the total time on target during the 5 min. session. Another was derived from the curve relating hits and total time on target for each subject, being the number of hits at the maximum of the curve which was at 55 per cent time on target. This curve was derived by fitting a parabola to the 10 points given by the relation between total time on target and hits in each 30 sec. period of the 5 min. session. This procedure is exactly the same as the one used in the previous experiment (Frith, 1968) except that each curve is based on only 10 points from 30 sec. periods, whereas previously it was based on 30 points from 10 sec. periods.

Another measure was derived to indicate the frequency of occurrence of rest pauses. This was simply the number of miss lengths greater than 1000 msec. This cut-off point was chosen quite arbitrarily, the measure being supposed to provide information about particularly long misses.

The last two measures were concerned with the occurrence of rhythmicities in performance. The repetitive nature of the task and the shape of the track lead one to expect that rhythmicities should be present in performance. This would be especially likely if the subjects attempted to match their velocity with that of the target, since such subjects would try to make rhythmic movements. Thus it was hoped that the amount of rhythmicity present would indicate whether the subjects were paying attention to velocity errors rather than position errors. Such rhythmicities were investigated by calculating the spectral density function for pursuit rotor performance. This function was estimated with two LINC-8 programs written by the author. The first autocorrelated 30 sec. of pursuit-rotor performance and the second converted these autocorrelations into the spectral density function via a cosine transformation (Robinson, 1967). The data were 30 sec. taken from the middle of the 5 min. working session, i.e. from 2·5 to 3·0 min. For most subjects two peaks appeared in this function at 1·60 sec. per cycle and at 0·53 sec. per cycle, which

corresponded to the time for a complete revolution of the target and the time for the target to move down one side of the triangular track. The value of the spectral density function at each of these points for each subject was the last two measures used to describe performance.

Table 1 gives a list of all measures and their derivations.

Table 1. *Measures used to describe pursuit rotor performance*

1. Total time on target (TTT) in a 5 min. session.
2. Probability of a hit starting after 150 msec.
   $$P(h) \geqslant 150 = \frac{\text{no. of misses between 150–350 msec.}}{\text{no. of misses} \geqslant 150 \text{ msec.}}$$
3. Probability of a miss starting after 150 msec.
   $$P(m) \geqslant 150 = \frac{\text{no. of hits between 150–350 msec.}}{\text{no. of hits} \geqslant 150 \text{ msec.}}$$
4. Probability of a hit starting before 150 msec.
   $$P(h) < 150 = \frac{\text{no. of misses} < 150}{\text{total no. of misses}}$$
5. Probability of a miss starting before 150 msec.
   $$P(m) < 150 = \frac{\text{no. of hits} < 150 \text{ msec.}}{\text{total no. of hits.}}$$
6. Degree of bimodality of hit distribution.
   Bimod. = (no. of hits between 400 and 450 msec.)
   − (no. of hits between 300 and 350 msec.)
7. Estimated number of hits at 55 per cent total time on target.
   h_{max} is estimated from the best fit parabola relating hits and TTT in each 30 sec. period.
8. Rhythmicity with a period of 1600 msec. (1 rev.) from minutes $2\frac{1}{2}$–3 of the session.
9. Rhythmicity with a period of 533 msec. (1 side) from minutes $2\frac{1}{2}$–3 of the session
10. Number of misses $\geqslant$ 1000 msec.

Intercorrelations between the measures

Table 2 shows the correlation matrix for the 10 variables described and also the two personality measures of extraversion and neuroticism. In a principal component analysis of this matrix the first two factors accounted for about 60 per cent of the variance. Rotation in the plane defined by these two factors enabled one to define two alternative factors with about equal share of the variance, one of which loaded very highly on level of attainment (i.e. total time on target). The other variables loading most highly on this factor were $P(h) \geqslant 150$, $P(m) \geqslant 150$, and misses $\geqslant 1000$.

There is nothing surprising in this cluster of related measures. As expected, the probability of hits and misses starting related very strongly to total time on target. This gives some confirmation for the applicability of the Poisson model and shows that these measures based on the frequencies of hit and miss lengths from 150–350 msec. reflect a fundamental variable. It was also to be expected that subjects with large numbers of long misses should have a low level of attainment.

The second factor was chosen to be independent of total time on target and can therefore be represented by the correlation matrix in Table 3, from which total time on target has been partialled out. This table, since it shows relations independent of level of attainment, defines the strategies used by the subjects. The table reveals a

large cluster of related variables centred on the probabilities of hits and misses starting. Before total time on target was partialled out these correlated negatively (−0·542) but afterwards they correlated positively (0·583). The other variables loading most highly on this factor of strategy were h_{max}, rhythmicity (one rev.), bimodality and $P(h) < 150$.

Table 2. *Intercorrelations between measures of pursuit rotor performance for 30 subjects*

	TTT	Bimod.	$P(m)$ ≥ 150	$P(h)$ ≥ 150	$P(m)$ < 150	$P(h)$ < 150	h_{max}	One rev.	One side	Miss ≥ 1000	E
TTT	.										
Bimod.	0·07	.									
$P(m)$ ≥ 150	−0·86	−0·22	.								
$P(h)$ ≥ 150	0·81	−0·14	−0·54	.							
$P(m)$ < 150	−0·37	−0·12	0·45	−0·18	.						
$P(h)$ < 150	0·86	0·46	−0·83	0·55	−0·36	.					
h_{max}	0·13	0·08	0·23	0·42	0·25	0·25	.				
One rev.	−0·33	−0·01	0·13	−0·54	0·09	−0·32	−0·30	.			
One side	0·05	−0·03	−0·02	0·22	−0·26	−0·17	−0·06	−0·35	.		
Miss ≥ 1000	0·88	−0·15	0·67	−0·76	0·25	−0·73	−0·19	0·35	0·03	.	
E	0·14	0·45	−0·39	−0·15	−0·12	0·33	−0·32	0·01	0·18	−0·07	.
N	−0·07	−0·16	0·07	−0·25	0·03	−0·00	−0·19	0·00	−0·26	0·16	−0·11

Table 3. *Intercorrelations between the measures of pursuit rotor performance, attainment level (TTT) partialled out*

	Bimod.	$P(m)$ ≥ 150	$P(h)$ ≥ 150	$P(m)$ < 150	$P(h)$ < 150	h_{max}	One rev.	One side	Miss ≥ 1000	E	N
Bimod.											
$P(m)$ ≥ 150	−0·31										
$P(h)$ ≥ 150	−0·36	0·58									
$P(m)$ < 150	−0·07	0·28	0·21								
$P(h)$ < 150	0·68	−0·33	−0·51	−0·07							
h_{max}	0·07	0·69	0·54	0·31	0·02						
One rev.	0·01	−0·34	−0·50	−0·02	−0·10	−0·28					
One side	−0·03	−0·13	0·47	−0·30	−0·24	−0·05	−0·39				
Miss ≥ 1000	−0·18	−0·38	−0·18	−0·16	0·13	−0·16	0·13	−0·01			
E	0·46	−0·54	−0·48	−0·08	0·45	−0·35	0·07	−0·17	0·12		
N	−0·15	0·01	−0·33	0·00	0·13	−0·19	−0·10	−0·24	0·20	−0·10	

It can be seen from the partial correlation matrix (Table 3) that extraversion score also correlates highly with many of these measures. It is now necessary to consider the interpretation of this cluster of measures in terms of actual behaviour.

At least one of these relationships is a necessary result of the definitions of the measures. The finding that when total time on target is partialled out there is a positive correlation between the probability of a miss starting and the probability of a hit starting merely reflects the fact that for a given total time on target a subject who is more likely to come off target must be more likely to come on target (thus producing many short hits and misses). This aspect of performance is also measured by the variable h_{max} (the number of hits at 55 per cent time on target) which would therefore be expected to form part of the cluster. This measure was the one used to define strategy in the previous experiment and thus it seems that the cluster of variables defines the same strategy as was found in that experiment. However, with the

additional measures taken it should now be possible to define these strategies more completely. It is not of course entirely necessary that $P(m)$ and $P(h)$ should be so closely related to the strategy defined in the first experiment since these measures are derived from hit and miss lengths between 150 and 350 msec. only. Thus the fact that these measures do relate so closely suggests that the strategies do concern these particular major sections of the hit and miss distributions and are not concerned merely with the occurrence of abnormally long misses or rest pauses. This is also suggested by the finding that the measure concerned with the frequency of long misses does not show either consistent or significant correlations with the various measures defining strategy.

There are three new and as yet unexplained aspects of strategy, which are bimodality, probability of hits starting in less than 150 msec., and rhythmicity with a period of one revolution of the target. The first two of these measures seem to be particularly highly related to each other (correlation 0·68). The probability of short misses loads in the opposite direction to the other hit and miss measures so that people characterized by long and therefore few hits and misses in the region above 150 msec. tend to show many of the very short misses.

Discussion

It will be demonstrated later that there is an explanation for these effects by considering the particular tracking task used in greater detail. But first some remarks can be made concerning the rhythmicity measure. This measure is such that people showing in general relatively long and therefore few hits and misses (the low hit strategy of the previous experiment) show a high degree of rhythmicity with a period of one revolution of the target. It seems likely that this rhythmicity reflects concern with errors of velocity-matching. If a subject has matched velocity and is on target he will do very well. However, when he has matched velocity and is off target he will do very badly. Thus one might expect that his performance would be characterized by relatively long hits and misses resulting in the relation between measures that was found in the present study. A person who is concerned with position errors will correct such errors very quickly, thus showing short misses, but will fail to stay on target for long since he has not accurately matched velocity. Thus such a subject will show the high hit strategy of the previous experiment.

This detailed analysis of pursuit rotor performance has confirmed the results of the previous experiment demonstrating a strategy of pursuit rotor performance which is independent of level of attainment, and which relates significantly to the extraversion dimension of personality. However, the hypothesis that this dimension of individual differences would be related to the occurrence of rest pauses does not seem to be confirmed, since the dimension is concerned particularly with hit and miss lengths of 350 msec. and less. On the basis of the measures taken it is tentatively hypothesized that the differences in strategy reflect the subjects' concern with either errors of velocity or errors of position.

An estimate was made of the reliability of the two principal components of performance. Level of attainment (total time on target) showed a test–retest reliability of 0·700 (comparison of second and third 5 min. work sessions). A suitable measure of strategy was devised by adding together the principal measures that defined this

dimension, i.e. Ph 150–350 and Pm 150–350. Since these two measures correlate equally with total time on target but in opposite directions, adding them together effectively removes any relation to level of attainment. The correlation between this derived measure of strategy and total time on target was 0·032. The test–retest reliability of this strategy measure was 0·688. Thus both measures seem to be stable and characteristic of an individual's performance.

So far no attention has been paid to the particular nature of the task used in this experiment. The target that the subjects had to follow moved round a triangular track, and therefore at the corners of the triangle the target abruptly changed direction. In addition, the target was produced by a radial band of light revolving underneath the track. It therefore moved much faster at the corners since these are farthest from the centre of revolution. Thus there were two factors making the target much more difficult to track at the corners of the triangle. It might be expected that these variations in task difficulty would affect the shape of the hit and miss distributions and might account for the bimodality shown by some subjects in the hit distribution. One noticeable feature of these distributions was that their modes occurred at the same points independent of both level of attainment and strategy. This suggests that the position of the modes is a function of the task itself. The target was revolving at 37·5 r.p.m. and so one revolution corresponded to 1600 msec. and one side of the triangle to 533·3 msec. Thus the two peaks of the bimodal hit distributions (250 and 450 msec.) corresponded to just under half a side and just under a side. Thus bimodality could be accounted for if the subjects' were particularly likely to come off in the middle of a side and at the end of a side. As has already been shown, subjects are particularly likely to come off at the end of a side when the corner occurs, but it seems very unlikely that they should come off at the middle of a side since at this point the target is moving more slowly than at any other time and in a straight line.

In order to investigate further the bimodality of the hit distribution two subjects were chosen from the original group who, while still available for testing, showed respectively extreme bimodality and extreme lack of bimodality. These subjects worked for another 5 min. with the triangular pursuit rotor, their performance on the last 2½ min. being filmed with an 8 mm cine camera. When these films were viewed, differences in the subjects' performance were immediately apparent. There were two aspects of these differences. The first lay in the path followed by the subject's tracking stylus. Figs. 2 and 3 show the position of the end of the stylus in relation to the triangular track on successive frames of the film. The subject with the unimodal hit distribution (Fig. 2) shows a triangular course closely related to the track. The subject with the bimodal hit distribution (Fig. 3) follows a roughly circular course which thus necessarily cuts the corners of the triangle and sometimes swings out from the centre of the sides. The second aspect of the differences lay in the relationship between the end of the subject's tracking stylus and the position of the target. The subject whose stylus followed a circular course was nearly always radially in line with the target, even when he was not actually on the track. The subject following the triangular course, although nearly always on the track, was often in front of or behind the target. He showed a strong tendency to be behind the target immediately after rounding a corner and then to catch up with it rapidly to the extent of overshooting, especially since in terms of linear velocity the target was at this stage slowing down. Thus it is

lear, due perhaps to the paths the subjects choose to trace out, that one of them is concerned principally to match the angular velocity of the target, while the other is concerned with the spatial location of the target.

It is clear also that a bimodality of the hit distribution can result from the subject following a roughly circular path with his stylus. With such a path the subject will not only cut the corners of the triangle, but also will sometimes swing too far out at the centres of the sides. When this happens he will produce two hits with lengths equal to somewhat under half a side each. On those occasions when he does not swing too far out at the centre of a side he will produce one hit of approximately double the

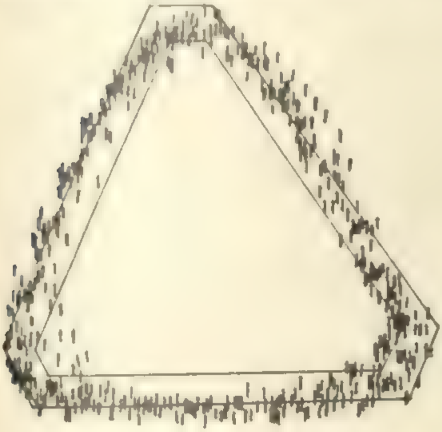


Fig. 2

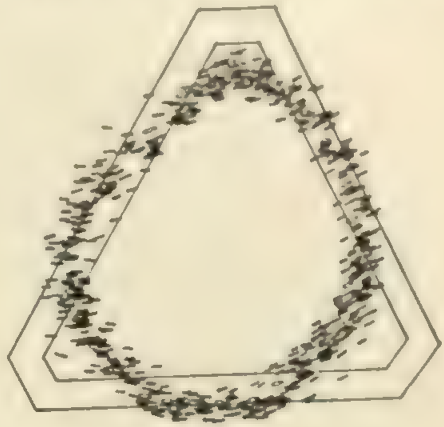


Fig. 3

Fig. 2. Successive positions of the tracking stylus for a subject with a unimodal hit-length distribution.

Fig. 3. Successive positions of the tracking stylus for a subject with a bimodal hit-length distribution.

length. Intermediate values between these two will be much less frequent. This strategy of matching velocity will produce particularly short misses, since the subject, when returning to the track after cutting a corner or swinging too far out from the centre of a side, is automatically anticipating the position of the target, rather than having to catch up with it or take time for correcting movements. This will lead to a greater proportion of short misses and thus accounts for the high correlation between the probability of a hit starting in less than 150 msec. and bimodality.

This explanation of the bimodality of the hit distribution makes it fit in very well with the other measures of performance that were found to relate to bimodality, since these also suggested a dimension of strategy concerned with either velocity- or position-matching. However, it also shows that the shape of the track plays an important role in the determination of strategy. One way in which the photo-electric pursuit rotor differs from the one previously used is that the track the target follows is visible at all times, whereas with standard pursuit rotors only the target is visible. If, with the photo-electric pursuit rotor, subjects thought it an important part of the task to keep their stylus at all times in the right track even when not on target, they would be forced to adopt the position-matching strategy described above. It is entirely

consistent with the characteristics imputed to extraverts (risk-taking, not conscientious, etc.) who showed a significant tendency to bimodality that they should cut the corners of the triangle rather than stay rigidly in the tracks. Thus it might be that the differences between introverts and extraverts found in these experiments are specific to this kind of task in which the track is visible and not related to a more general favouring of velocity or position-matching.

It is clearly crucial to the hypothesis being put forward to study different track shapes and target velocities, since these variables should also influence the strategies adopted and their relation to personality. With regard to track shape it would be predicted that the farther removed this were from a circle in shape the more position- rather than velocity-matching would be adopted. Some evidence for this is provided by an earlier experiment (Frith, 1968) in which performance on the triangular track of the present study was compared with performance in a star-shaped track. Strategies of performance in terms of long or short hits were measured which in the light of the present results would be interpreted as velocity- and position-matching respectively. For both track shapes there was a significant relationship between strategy and personality, with extraverts tending to adopt velocity-matching. However, there was also a significant difference between tracks in the effects of practice on strategy. While practice on the star-shaped track had no effect on strategy, there was a general shift towards velocity-matching for the triangular track ($F = 26{\cdot}15$; d.f. $= 1, 18$; $P < 0{\cdot}001$) which was significantly greater for the introverted subjects ($F = 4{\cdot}56$; d.f. $= 1, 18$; $P < 0{\cdot}05$). These results support the hypothesis that the simpler triangular track induced velocity-matching and also that the introverts, since they tended to adopt position-matching, were more affected by this induction.

With regard to target speed it would be predicted that higher speeds would tend to induce velocity-matching since the fundamentally circular movement of the target would be more apparent and the position of the target at any instant would be more difficult to estimate. Performance of subjects in the present experiment has been compared with that of people practising on the same triangular track, but with a target speed of 50 r.p.m. instead of 37·5 r.p.m. I am indebted to Dr Gudrun Sartory for allowing me to use her data for this comparison. When groups were compared who had the same range of total time on target scores, the subjects tracking a target at 50 r.p.m. were significantly more likely to have long-hit strategies ($\chi^2 = 21{\cdot}08$; $P < 0{\cdot}001$). If long-hit strategies can be taken to indicate velocity-matching then this result confirms the hypothesis that high target speeds induce velocity-matching. For the group with the high target speed there was no relationship between strategy and extraversion. This would suggest that at this high speed all subjects rapidly adopt the strategy of matching velocity.

This experiment has confirmed the hypothesis that there are differences in the style of pursuit rotor performance between introverts and extraverts. However, this difference did not appear to result in extraverts producing more rest pauses than introverts. The cluster of measures defining the strategies rather suggested that the extraverts principally concerned themselves with errors of velocity-matching while the introverts principally concerned themselves with errors of position-matching. It is hoped that studies of pursuit rotor performance in terms of these strategies may be able to throw light on other phenomena such as the greater reminiscence shown by

extraverts as compared to introverts. It has been proposed that this difference between introverts and extraverts results from the greater proneness of the extraverts to the build-up of inhibition (Eysenck, 1956). This hypothesis could be reconciled with the present results if it is assumed that this inhibition build-up results, not in the production of rest pauses, but in a generally slower response rate. It is assumed that the responses in pursuit rotor performance consist of the detection and correction of errors and are therefore intermittent (Poulton, 1957). It seems probable that a slow rate of responding would favour velocity-matching whereas a fast rate of responding would favour position-matching. Estimation of velocity would require the integration of several successive positions of the target and would therefore take longer to estimate than position.

This study was supported by a grant from the Bethlem Royal and Maudsley Hospitals Research Fund. I am grateful to Professor H. J. Eysenck for his help and encouragement.

References

Eysenck, H. J. (1956). Reminiscence, drive and personality theory. *J. abnorm. soc. Psychol.* **53**, 328–333.

Eysenck, H. J. (1965). A three-factor theory of reminiscence. *Br. J. Psychol.* **56**, 163–181.

Eysenck, H. J. & Eysenck, S. B. G. (1964). *Eysenck Personality Inventory*. London: University of London Press.

Frith, C. D. (1968). Strategies in rotary pursuit tracking and their relation to inhibition and personality. *Life Sci.* **7**, 65–76.

Hull, C. L. (1943). *Principles of Behaviour*. New York: Appleton–Century–Crofts.

Kimble, G. A. (1949). An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.* **39**, 15–23.

Poulton, E. C. (1957). On the stimulus and response in pursuit tracking. *J. exp. Psychol.* **53**, 189–194.

Robinson, E. A. (1967). *Multichannel Time Series Analysis with Digital Computer Programs*. London: Holden-Day.

Spielman, I. (1963). The relation between personality and the frequency and duration of rest pauses during massed practice. (Unpublished Ph.D. thesis, University of London Library.)

(*Manuscript received* 17 *April* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 199–215
Printed in Great Britain

REMINISCENCE AND THE SHAPE OF THE LEARNING CURVE AS A FUNCTION OF SUBJECTS' ABILITY LEVEL ON THE PURSUIT ROTOR

By H. J. EYSENCK and J. E. GRAY

Department of Psychology, Institute of Psychiatry, University of London

Two experiments are reported in which the shape of the learning curve pre-rest and post-rest, and the reminiscence phenomenon, are studied as a function of the ability of the subjects to perform on the pursuit rotor under conditions of massed practice. The experiments differ in the choice of subjects and the choice of practice and rest periods involved. Both experiments demonstrate unequivocal differences in the performance curves of high- and low-ability subjects, both before and after the rest pause; these differences are in part also a function of the length of the rest pause. Reminiscence was also found to depend on ability level, high reminiscence being found in high-ability subjects. The applicability of several different hypotheses to the phenomena in question is discussed and a theory suggested which combines features from several earlier theories.

Some of the parameters determining reminiscence on the pursuit rotor are reasonably well known; examples are length of massed pre-rest practice, length of rest pause, and level of motivation (Eysenck, 1964*a*). Much less is known about other phenomena closely associated with reminiscence, e.g. post-rest upswing (PRU), also sometimes called warm-up decrement (Ammons, 1947) and post-rest downswing (PRD). There are also very marked individual differences in the occurrence of these phenomena; these usually appear in the analyses in the form of unduly large error variances. Some of these individual differences have been shown to be associated with personality variables such as extraversion–introversion (Eysenck, 1962; Farley, 1966; Gray, 1968), but most of the variance is still unaccounted for. Under these circumstances it seemed worth while to look at a variable which produces tremendous differences between subjects, namely their level of ability. From the very beginning of practice some subjects perform at a level several hundred or even thousand per cent above that of other subjects of similar background and intelligence, and these differences tend to persist over hours of continued practice (Jones, 1966), whether massed or distributed. The possibility that ability differences may affect the shape of the learning curve cannot be dismissed without close examination, and we have attempted in this paper to demonstrate certain phenomena associated with high and low ability (defined in terms of initial performance) respectively.

There is little information in the literature on this point, and what there is is somewhat contradictory (Buxton & Grant, 1939; Leavitt, 1945; Leavitt & Schlosberg, 1944; Reynolds & Adams, 1954; Zeaman & Kaufman, 1955; Cieutat & Noble, 1958; Locke, 1965; Jahnke, 1961; Eysenck, 1964*b*; Clark, 1967). Some of these studies are only marginally relevant, and in only few of them has the problem been attacked directly. Reynolds & Adams (1954) report the only major study in this field, comparing subjects trained 'one group of subjects with massed and a second group with distributed practice' (p. 269) and subdivided according to initial level of ability. They concluded that 'with the exception of slope characteristics of first-session curves

no evidence has been found for the interaction of ability-level and learning variables' (p. 276). This finding, however, is suspect in so far as the 'massed' practice is concerned; it seems that each practice period of 20 sec. was followed by a rest period of 5 sec. in the so-called 'massed' practice, thus effectively converting it into a 'spaced' type of practice. Eysenck (1964*b*) reported data on 300 high-drive apprentices, divided into five ability groups. Results showed (1) that in the pre-rest practice session the high-ability group showed a marked initial upswing which after 2 min. turned

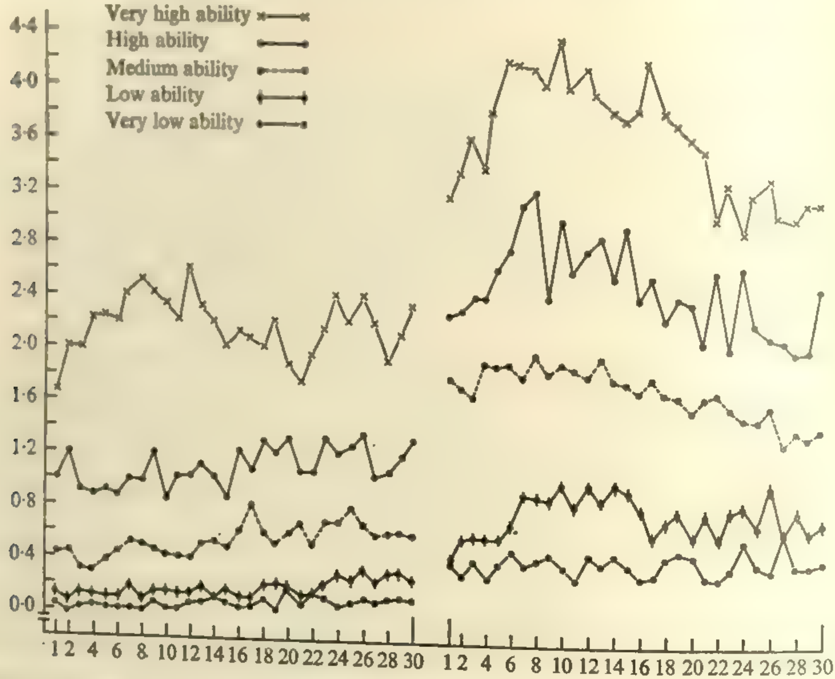


Fig. 1. Pre-rest and post-rest performance of different ability groups on the pursuit rotor. A 10 min. rest pause is interpolated between practice periods.

into a downswing. (2) Reminiscence scores did not differ significantly between ability groups. (3) PRU and PRD were clearly related to ability level, being stronger in the higher scoring groups, and in fact inverted in the lowest scoring group, where post-rest upswing is absent and a marked post-rest downswing present instead (Eysenck, 1964*a*, fig. 2). It should be noted that these data are derived from a high-drive group, in which motivation was produced in the manner described by Eysenck (1964*a*); it does not follow that low-drive groups would show similar patterns of reaction.

Clark (1967), in an unpublished study, repeated Eysenck's study on five pursuit-rotor ability groups, each of 25 subjects, defined by rank ordered performance level of the first 10 massed practice levels; subjects were hospitalized schizophrenics. Fig. 1 shows his main results, which are very similar to those reported by Eysenck (1964*b*); note the initial upswing of the high-ability group during pre-rest practice, and the PRU and PRD characteristically diminishing from high- to low-ability groups. Neither age nor duration of illness had any association with ability on the rotor. Unlike the

Eysenck study, however, there were very marked differences in reminiscence between the groups, which by analysis of variance exceeded a $P = 0{\cdot}001$ level: 'mean reminiscence scores of the very high, high, and medium ability groups are significantly greater than the mean scores of each of the low and very low ability groups' (Clark, 1967, p. 182). This effect may be an artifact due to the very low level of performance of the worst two groups; with normal groups many of these subjects would have been rejected according to the rules imposed by Eysenck (1964*b*), namely that 'subjects who failed to learn the task were eliminated and others used to replace them, the criterion of "learning" being a score of at least 1 sec. on target during at least one of the 30 10-sec. periods which constituted the pre-rest practice period' (p. 180). However, the data may also be interpreted as suggesting that with low-drive groups there is a tendency for high-ability subjects to show more reminiscence than low-ability subjects. It was part of the purpose of the experiment to be reported to furnish further information on this point.

Experiment I

Two experiments were in fact performed, using rather different designs, in order to investigate various parametric determinants of the phenomena in question. The theoretical relevance of these parametric investigations will be discussed in a later section. In the first experiment, eight groups of low-drive industrial apprentices were given massed practice on the pursuit rotor for either 3 min. or 8 min., scored in terms of 10 sec. periods time-on-target; rest periods of either 30 sec., 2 min., 6 min. or 20 min. followed pre-rest practice, and the rest periods in turn were followed by another 4 min. of practice for all groups. Reminiscence scores were derived by subtracting the mean of the last three pre-rest trials from the first post-rest trial; using the mean of three trials makes the score more reliable, and inspection shows little systematic change in performance at this stage. A single post-rest trial has to be used because of PRU which produces marked and systematic changes in performance.

Each of the eight groups contained 30 subjects; of these the top 12 and the bottom 12 in ability were chosen for purposes of analysis. Ability was defined in terms of total performance over the first 12 trials, and analysis of variance showed no significant differences in performance between groups either for the high-ability or for the low-ability subjects; in other words, we can compare our two sets of performance groups without having to control for initial differences in level of ability within sets. No detailed description is here given of either the type of population or the instrument used; details regarding both have been given elsewhere (Eysenck, 1964*a*). The subjects were also administered the Maudsley Personality Inventory, but scores of this merely confirmed that allocation to groups had indeed been random, and that ability did not correlate with personality, at least in so far as this is measured by the inventory.

Pre-rest performances of the 96 subjects who practised for 3 min. are diagrammed in Fig. 2; those of the 96 subjects who practised for 8 min. are diagrammed in Fig. 3. Inspection suggests a simple linear increase in performance for the low-ability groups, whilst the high-ability group shows a marked upswing to begin with, followed by a plateau; for the 8 min. group there is no improvement from the 10th trial to the

48th. (The 3 min. group does not appear to have settled down sufficiently to make it possible to assess the applicability of the plateau notion.) Analysis by orthogonal polynomials was performed on the data for both groups, to give statistical background to the results of the visual inspection. For the 8 min. group the linear component was significant at the $P < 0{\cdot}001$ level for both high- and low-ability subjects; the quadratic component was also significant for both groups ($P < 0{\cdot}01$). Cubic and quartic components were insignificant in the low-ability group, but significant at the $P < 0{\cdot}01$ and $< 0{\cdot}05$ levels respectively for the high-ability group. Thus for all intents and purposes the low-ability group shows a linear increment with some evidence of a bend, while the high-ability group, in addition to the linear component,

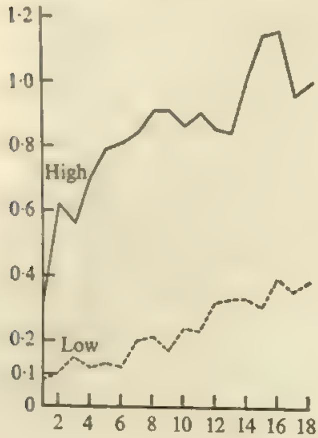


Fig. 2

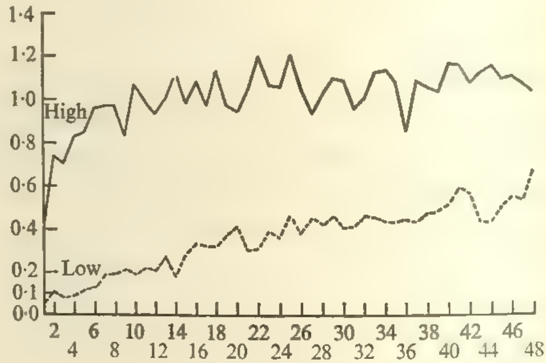


Fig. 3

Fig. 2. Pre-rest performance of high-ability and low-ability subjects during 3 min. practice period.

Fig. 3. Pre-rest performance of high-ability and low-ability subjects during 8 min. practice period.

shows higher-order variations up to the 4th power. Analysis of the combined high- *v.* low-ability groups show linear, cubic and quartic components to give rise to significant differences ($P < 0{\cdot}05$, 0·05 and 0·01 respectively). Analysis of the 3 min. groups shows a highly significant linear component for both high- and low-ability subjects ($P < 0{\cdot}001$); the only other significant value is the quadratic for the high-ability group ($P < 0{\cdot}01$). Analysis for the combined groups (high or low ability) confirms that it is in this component that the two groups differ most significantly ($P < 0{\cdot}01$); however, linear and cubic trends also show significance at the 0·05 level. Visual inspection therefore is largely vindicated by the analysis, except that the high-ability groups show a record even more complex than suggested, including cubic and quartic trends.

Post-rest data are given for the 30 sec. rest groups in Figs. 4 and 5; these show PRD but no sign of PRU. Furthermore, the high-ability group shows greater reminiscence than does the low-ability group. Analysis by orthogonal polynomials shows that the linear decrement is significant ($P < 0{\cdot}05$) for the high-ability group, but not for the low-ability group; no other powers are significant. Analysis for the combined groups shows them to be performing at significantly different levels

($P < 0.001$) with the linear element differentiating significantly between their rates of decline ($P < 0.05$).

Post-rest data are given for the 2 min. rest groups in Figs. 6 and 7; these show PRD (Fig. 6) but no sign of PRU. The high-ability group again shows greater reminiscence than does the low-ability group. Analysis by orthogonal polynomials fails to show statistical significance for any sequence effect, so that the visual appearance of decrement is not supported.

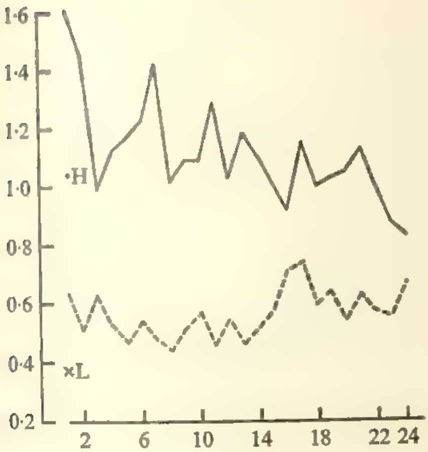


Fig. 4

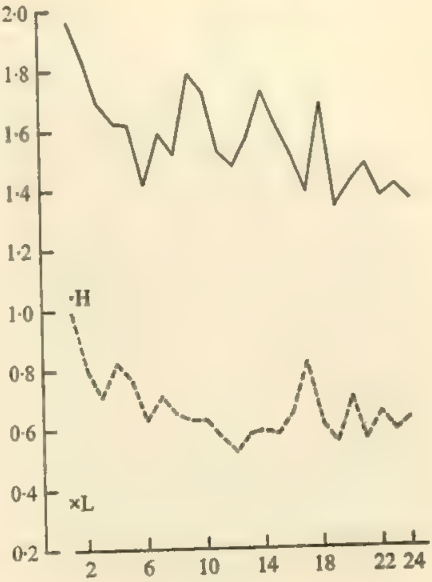


Fig. 5

Fig. 4. Post-rest performance for 30 sec. rest group, after 3 min. of pre-rest practice. (In Figs. 4–11 the pre-rest performances of the high-ability and low-ability groups are marked with an H or an L respectively to enable the reminiscence score to be estimated visually from the figure.)

Fig. 5. Post-rest performance for 30 sec. rest group, after 8 min. pre-rest practice.

Post-rest data are given for the 6 min. rest groups in Figs. 8 and 9. The high-ability group shows significant decrement for both the 8 min. practice group ($P < 0.01$) and the 3 min. practice group ($P < 0.001$). An analysis of the combined groups shows the decrement difference to be significant ($P < 0.05$) for the 3 min. group only. For the 8 min. group there is also a significant difference in the quartic ($P < 0.05$). Reminiscence scores are again higher for the high-ability group.

Post-rest data are given for the 20 min. rest groups in Figs. 10 and 11; these show both PRU and PRD. The high-ability group again shows greater reminiscence than does the low-ability group. Analysis by orthogonal polynomials shows both PRU and PRD to be present in the high-ability group with 3 min. practice, but not in the low-ability group; linear and quadratic components are significant for the former ($P < 0.05$), but not for the latter. No other components are significant. For the 8 min. groups only the quadratic component is significant in the high-ability subjects; neither component is significant for the low-ability group. Visual inspection suggests that for the high-ability group 8 min. of pre-rest practice produces a stronger PRU

which cancels in extent the PRD, resulting in a non-significant linear component; for the low-ability group neither PRU nor PRD are strong enough to produce statistically significant results, possibly due to the relatively small number of subjects in each group.

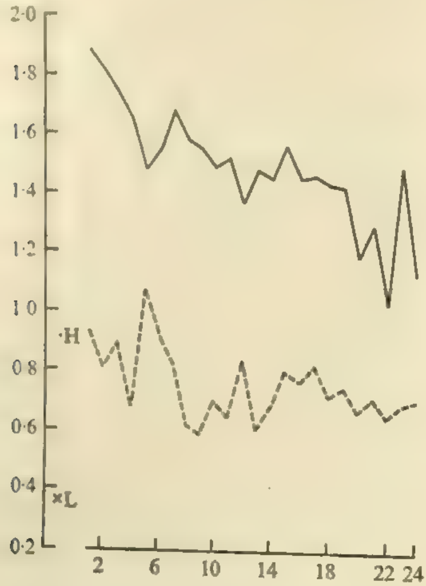


Fig. 6

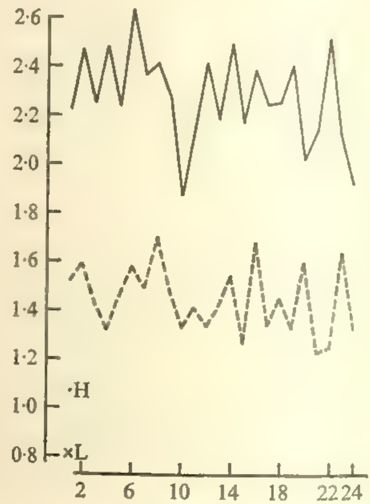


Fig. 7

Fig. 6. Post-rest performance for 2 min. rest group, after 3 min. of pre-rest practice.
Fig. 7. Post-rest performance for 2 min. rest group, after 8 min. of pre-rest practice.

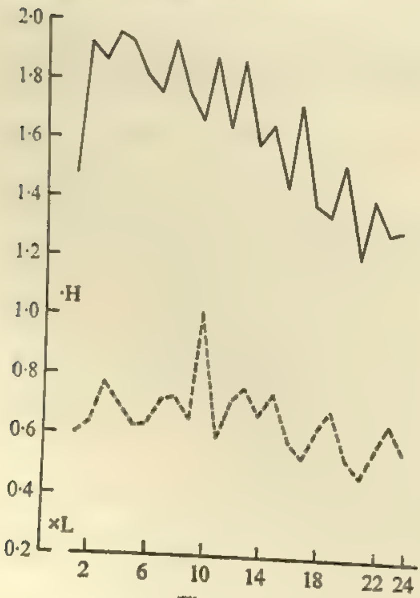


Fig. 8

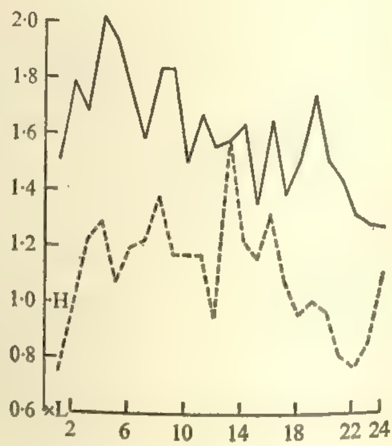


Fig. 9

Fig. 8. Post-rest performance for 6 min. rest group, after 3 min. of pre-rest practice.
Fig. 9. Post-rest performance for 6 min. rest group, after 8 min. of pre-rest practice.

In the combined high- and low-ability groups' analysis for 3 min. pre-rest practice, the linear component gives a significant differentiation ($P < 0{\cdot}05$), while for the combined 8 min. groups none of the polynomials results in significant differences. This suggests that subjects in the low-ability groups show a trend in the same direction as those in the high-ability groups, but much weaker; thus these trends are not themselves significant, but they are sufficiently strong to prevent differences between the groups from becoming significant.

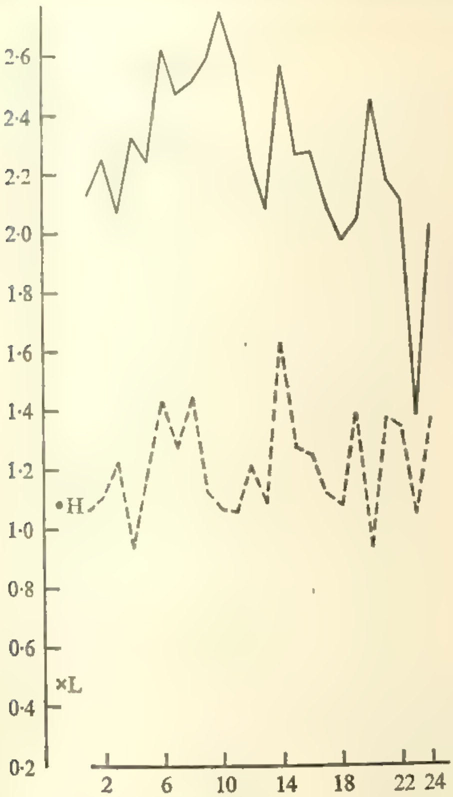


Fig. 10

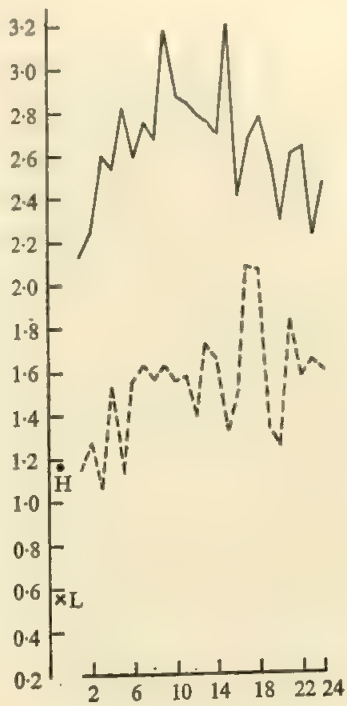


Fig. 11

Fig. 10. Post-rest performance for 20 min. rest group, after 3 min. of pre-rest practice.
Fig. 11. Post-rest performance for 20 min. rest group, after 8 min of pre-rest practice.

Differences in reminiscence scores have been mentioned, but not analysed so far. Analysis of variance was carried out on reminiscence as a function of high and low ability, long and short pre-rest practice, and length of rest period. High-ability groups in each case showed greater reminiscence; longer pre-rest practice produces greater reminiscence; and longer rest pauses produce greater reminiscence. Ability level produces statistical significance ($P < 0{\cdot}001$); rest reaches a significance level of $P < 0{\cdot}05$, while length of pre-rest practice was not significant. It is notable that ability level is a much more significant determiner of reminiscence than either of the other two factors in the genesis of reminiscence. None of the interactions were significant (cf. Buxton, 1943).

Results reported so far suggest the following conclusions. Pre-rest practice produces mainly linear trends in low ability subjects and complex non-linear trends in high ability subjects. Post-rest results show PRD, which is stronger in high ability subjects than in low ability subjects. Post-rest results show PRU only with long rest periods. PRU is more apparent in the high-ability groups than in the low ability groups. Reminiscence under all circumstances is greater in high-ability than in low-ability subjects. These results demonstrate that all parts of the learning curve are vitally affected by ability level, and suggest that simple averaging of combined data may simply confuse proper determination of learning curves.

EXPERIMENT II

In this experiment all subjects did 6 min. of massed practice, followed by a rest of 30 sec., 10 min. or 1 week; post-rest practice consisted of a further 7 min. of practice. Two concentric targets were used in order to study effects of ease or difficulty of the task (Bahrick *et al.*, 1957, have drawn attention to the fact that marked differences in the shape of learning curves may result from such differences.) The smaller target was 0·5 in. in diameter, the larger 1·61 in.; details of the apparatus are given in Gray (1968). Subjects in each rest group were divided into three subgroups of equal number on the basis of their cumulative time on target (0·5 in. target) over the first 120 sec. of practice. The Maudsley Personality Inventory was given to all subjects, but only served to confirm the random allocation of subjects to ability groups. There were approximately 70 subjects in each of the three different rest groups (72 in the 30 sec. group, 70 in the 10 min. group, and 68 in the 1-week group). Subjects were low-drive student volunteers, i.e. subjects who had no special and extraneous motivation to do well on the task.

Pre-rest scores of the three ability groups are shown in Figs. 12 and 13 for the difficult and easy targets respectively. Visually it seems that on the small target the high-ability group shows a distinct upswing followed by a distinct downswing, while the average ability group shows a straight but slow increase in performance, and the low-ability group an initial slight downswing followed by an increase in performance slightly faster than that of the average ability group. Analysis by orthogonal polynomials gave the following results. The high-ability group shows no significant linear trend, but highly significant quadratic and cubic trends ($P < 0{\cdot}001$). The average ability group has a very significant linear trend ($P < 0{\cdot}001$), and a significant quadratic trend ($P < 0{\cdot}01$), but no cubic. The low-ability group only has a linear trend ($P < 0{\cdot}05$). Combined analysis of all groups discloses, in addition to the obvious level difference ($P < 0{\cdot}001$), significant differences at the $P < 0{\cdot}01$ level for linear, quadratic and cubic trends.

For the easy target, the linear component is not significant; quadratic and cubic components are significant at the 0·001 level for the high-ability group. For the average ability group the linear component is highly significant, as is the quadratic ($P < 0{\cdot}001$); the cubic is barely significant ($P < 0{\cdot}05$). For the low-ability group the linear trend is highly significant ($P < 0{\cdot}001$), and the quadratic is barely so ($P < 0{\cdot}05$). The combined groups are differentiated, apart from the obvious differences in level ($P < 0{\cdot}001$), on the linear ($P < 0{\cdot}001$) and the cubic ($P < 0{\cdot}05$) trends. Taking

these results together with visual inspection of Fig. 12, we may say that the high-ability group shows marked upswing and downswing; the average-ability group less marked upswing followed by a plateau; and the low-ability group downswing followed by rapid increment. Change in target size may thus be said to have brought out and emphasized trends already present in the results of the difficult target; there are no marked changes in general shape of the learning curve. The initial decrement in the low-ability group, already suggested in the analysis of the results from the difficult

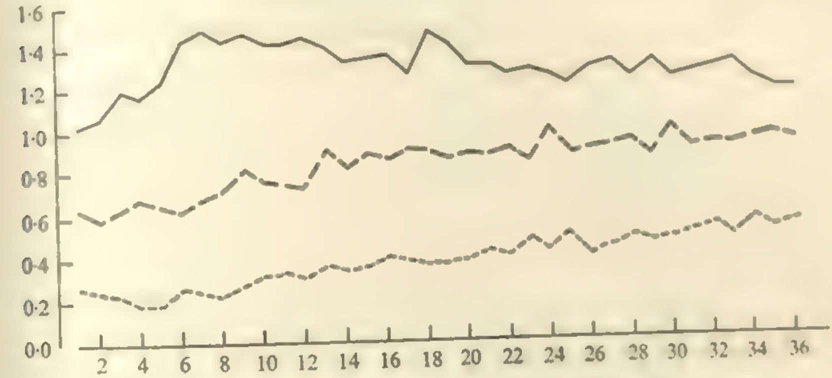


Fig. 12. Pre-rest scores of high-, medium- and low-ability groups, using the difficult target.

task, is clear and significant here; its causes are difficult to guess at and do not seem predictable from any existing theory. The high-ability group shows a learning curve surprisingly similar to that often found post-rest in massed practice groups: PRU and PRD are here in evidence as pre-rest upswing and pre-rest downswing. The possibility that both phenomena have the same or similar causes will be discussed later on.

Post-rest curves of practice are given below; Fig. 14 gives data for the 30 sec. rest groups, Fig. 15 for the 10 min. rest groups and Fig. 16 for the 1-week rest groups. Analysis by orthogonal polynomials was carried out, and the results for linear, quadratic and cubic components are shown in Table 1. It will be seen that length of rest pause is a very important factor. For the 30 sec. rest pause groups none of the results are significant; all groups proceed on a level without significant rises or decrements in performance. The 10 min. rest pause produces very marked effects, with linear, quadratic and cubic components all significant for all groups. For the combined groups only the linear trend is significant, apparently because of the PRD depressing the high-ability group's performance below its starting-point. For the 1-week group linear components are insignificant, quadratics very significant for the high-ability subjects, and insignificant for the low-ability subjects. Cubics are significant only for the high- and average-ability groups. On the combined groups analysis, both quadratic and cubic components are significant. These data leave no doubt that groups of different ability levels show highly significant post-rest performance curves provided the rest is long enough to permit consolidation (or dissipation of inhibition) to take place to a sufficient degree.

Reminiscence was calculated as before; the data show a significant interaction

effect. The low ability group shows a continued increment in reminiscence over the three rest periods; the other two groups show a rise from 30 sec. to 10 min. followed by a fall. These results are in good agreement with what one might have expected from Leavitt's findings: a positive correlation between ability and reminiscence for short rest pauses, and a negative one for long rest pauses.

A slightly more analytic way of presenting the data is the following. A linear regression coefficient '*b*' was computed for the final 4½ min. of the pre-rest practice period to indicate depression of performance during this time, i.e. following any initial upswing. A regression coefficient '*a*' was calculated to represent the level of ability measured 1½ min. after performance started. These coefficients were corre-

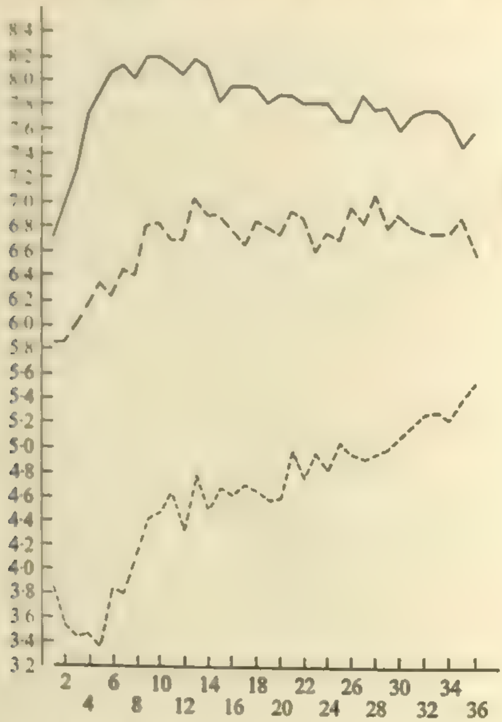


Fig. 13. Pre-rest scores of high-, medium- and low-ability groups, using the easy target.

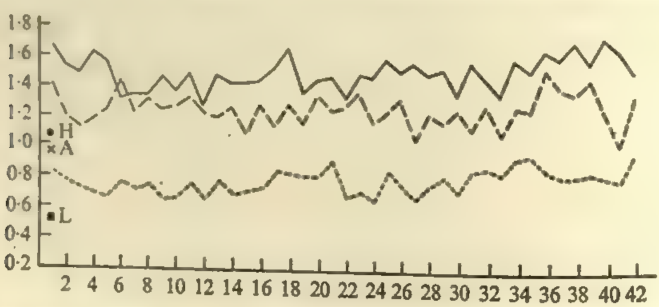


Fig. 14. Post-rest scores for high-, medium- and low-ability groups after 30 sec. rest pause.

ated with reminiscence scores, determined as before, and also with a new coefficient, called rem. max. This coefficient uses the same pre-rest measure of reminiscence, i.e. the mean of the final three pre-rest trials; however, the post-rest measure used is the mean of the two highest trials occurring at any time within the first 2 min. of post-rest practice. This index is designed to measure reminiscence to the top of each

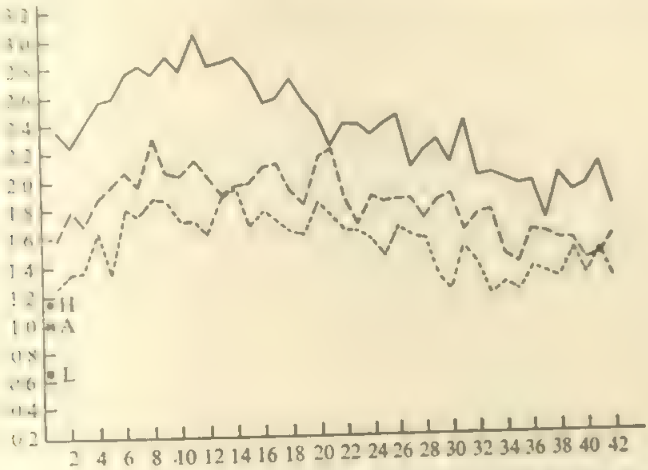


Fig. 15. Post-rest scores for high-, medium- and low-ability groups after 10 min. rest pause.

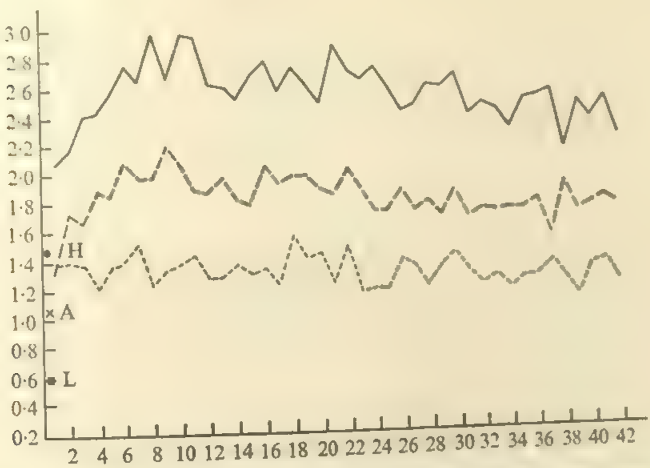


Fig. 16. Post-rest scores for high-, medium- and low-ability groups after 1-week rest pause.

subject's post-rest upswing; it has some resemblance to Ammons's (1947) index, but does not use his method of backward extrapolation. The correlations between reminiscence and rem. max. on the one hand, and '*a*' and '*b*' on the other, are given in Table 2; there is a gradual decrease in the correlations between reminiscence and '*a*' with increasing length of rest, but this change is not significant. Neither is the increasing correlation with '*b*', and the changes in size of correlation with rem. max.

are neither systematic nor significant. These data suggest that level of ability in this study is importantly correlated only with rem. max., i.e. an index which combines the twin effects of the slight correlation of reminiscence with ability with the strong PRU characteristic of high ability subjects.

Table 1 *Significance of linear, quadratic and cubic orthogonal polynomials for groups of high, average and low ability, and all groups combined, when tested with rest pauses of 30 sec., 10 min. or 1 week respectively*

	30 sec.	10 min.	1 week	
High ability	n.s.	0·001	n.s.	Linear
	n.s.	0·001	0·01	Quadratic
	n.s.	0·001	0·01	Cubic
Average ability	n.s.	0·01	n.s.	Linear
	n.s.	0·001	0·05	Quadratic
	n.s.	0·05	0·001	Cubic
Low ability	n.s.	0·01	n.s.	Linear
	n.s.	0·001	n.s.	Quadratic
	n.s.	0·001	n.s.	Cubic
Combined	n.s.	0·05	n.s.	Linear
	n.s.	n.s.	0·05	Quadratic
	n.s.	n.s.	0·05	Cubic

Table 2 *Correlations between reminiscence and 'maximum reminiscence' scores and regression coefficients 'a' and 'b', representing pre-rest performance level and decrement*

		Reminiscence	Rem. max.
30 sec. group	'*a*'	0·153	0·469**
	'*b*'	−0·261**	−0·239*
10 min. group	'*a*'	0·082	0·338**
	'*b*'	−0·210	−0·265*
1 week group	'*a*'	0·032	0·413**
	'*b*'	−0·416**	−0·245*

* $P < 0{\cdot}05$. ** $P < 0{\cdot}01$.

Results from this study, using older, brighter and possibly more highly motivated subjects than the previous one, confirm in most respects the conclusions already reached. Pre-rest performance shows high ability giving rise to upswing followed by downswing, while low ability shows the opposite pattern. Post-rest performance shows PRU and PRD to be phenomena peculiar to subjects in the higher-ability groups, and missing in the lower-ability groups; length of rest pause was also found to be an important factor in this connexion. Reminiscence was not found to be determined by ability level to anything like the extent suggested by the first experiment, or that of Clark (1967); it is possible that this may be due to greater motivation, making this group more comparable to Eysenck's (1964*b*) high-drive group, where no relation was found between ability and reminiscence. The apprentice group used in the first experiment was on the whole rather poorly motivated if one may use personal judgement based on observation, whereas the university students used in the present experiment seemed to be more ego-involved (Alper, 1948). Clearly motivation is an important variable in untangling the relationship between

ability and reminiscence, and equally clearly our data do not provide the required independent measure of motivation without which these suggestions cannot be regarded as anything but speculation.

Experiment III

In the Reynolds & Adams (1954) study it seems possible that their failure to observe post-rest differences between their different ability groups was due to the fact that they used distributed rather than massed practice. If this were true, then it should be possible to compare groups of low-drive apprentices, similar to those used in Expt. I, engaged on distributed learning of pursuit rotor; if distribution of practice is responsible for the failure of ability differences to mark differences in the shape of the learning curve, then such an experiment should result in essentially similar curves, excepting of course the course of learning preceding the first imposed rest.

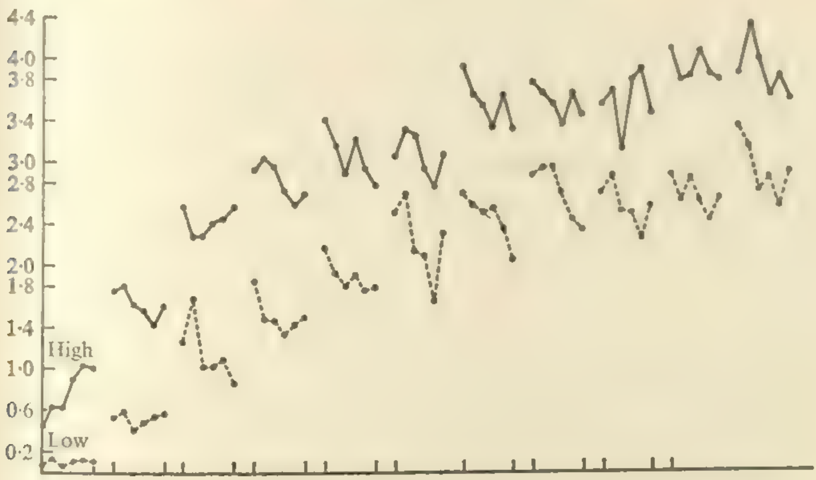


Fig. 17. Scores of high- and low-ability groups during 11 1 min. practice periods, separated by 5 min. rest pauses.

Fifteen high-ability and 15 low-ability subjects were chosen from 45 subjects who had practised on the pursuit rotor for 11 1 min. periods, separated by 5 min. rest pauses. Fig. 17 shows the results. It will be seen that PRU is universally missing in both groups, as would be expected from Eysenck's (1956) theory, on the basis that not enough practice time had been allowed for the accumulation of I_R and $_SI_R$. PRD's are observed equally in both groups; in Eysenck's (1965) theory this could be due to consolidation effects depressing performance post-rest. The only trial on which the groups differ significantly (by orthogonal polynomial analysis) is the first; as previously observed, the high-ability group shows a rapid upswing while the low-ability group shows no change at all. (A medium-ability group, formed from the remaining 15 subjects but not shown in the figure because it interferes with the

clarity of presentation without adding anything to the argument, showed intermediate upswing followed by a plateau.) These results support the view that massing is responsible for the major differences between high-ability and low-ability groups.

Discussion

The term 'ability' has hitherto been used as simply a descriptive label referring to the initial performance level of the subjects; as it happens initial performance levels tend to be preserved throughout the course of experiments like this, correlating reasonably highly with terminal performance levels, so that 'ability' so defined characterizes an individual's performance throughout the experiment. But clearly the effects of ability level on the shape of the learning curve cannot be discussed without some more analytic dissection of this concept of ability, which may be looked upon as either an innate property of the organism or an acquired skill, or a combination of both. Twin studies by McNemar (1933) demonstrated a heritability coefficient of 0·9 for pursuit-rotor performance (decreasing with increasing practice), and Vandenberg (1962) reported somewhat lower but still significant values. Environmental influences are less well documented, but it does not seem unreasonable to postulate that practice in such tasks as watching gramophone records go round, writing, drawing and tracing with pen or pencil, and turning the steering wheel of a motor car may produce practice increments in those fundamental abilities which underlie pursuit-rotor practice (Fleishman, 1960). It seems reasonable to assume that the phenotype of pursuit-rotor skill at the beginning of practice is in part a reflexion of genuine genotype differences, and in part the outcome of specific environmental histories. Can either or both of these components help us understand the phenomena connected with individual differences in (phenotypic) ability?

Reynolds & Adams (1954) suggest an interesting hypothesis. Drawing attention to the similarities between the pre-rest performance curves of high-ability subjects and the well-known PRU phenomenon, which they would explain in terms of warm-up or the regaining of some physical or mental set, they say: 'If the warm-up period can be regarded as a period of recruitment of previously acquired responses, then subjects in massed decile 10 (i.e. their highest ability group) would appear to be activating a pool of previously acquired relevant responses, carried over, perhaps, from psychomotor tasks encountered in everyday situations' (p. 276). The similarity on which they base this suggestion is even closer in our experiments, as we find PRD as well as PRU effects in our pre-rest practice of high-ability subjects, due no doubt to the fact that we used proper massing of trials, whereas Reynolds & Adams did not.

This additional feature, however, rather spoils the appeal of their theory as nothing in warm-up theory accounts for PRD, or for the downswing found here pre-rest. Furthermore, there are many arguments against the application of warm-up theory in relation to pursuit-rotor learning (Feldman, 1963) which apply equally cogently to pre-rest practice as to post-rest practice. The alternative hypothesis suggested by Eysenck (1956) for PRU, i.e. the extinction of ${}_{S}I_{R}$ through non-reinforcement, might be thought to fit the case rather better as all that is required is the assumption that higher-ability subjects have in the past practised components of the task more assiduously (massed practice) than low-ability subjects. Such an explanation, how-

ever, would leave out entirely the possibility of genetic differences between high- and low-ability subjects, and would appeal, as does Reynolds & Adams's explanation, entirely to unobserved and speculative events in the past lives of the subjects.

Both the 'warm-up' or 'set' theory and the extinction of $_sI_R$ hypothesis fail to account for the non-existence of PRU in the case of the low-ability subjects. It would need some special *ad hoc* assumptions to explain why low-ability subjects do not show 'warm-up' while high-ability subjects do; there is nothing in the literature to suggest anything of the kind (Ammons, 1947; Adams, 1961). The conditioned inhibition hypothesis might argue that high-ability subjects have worked harder than low-ability subjects during the pre-rest practice period, have accordingly accumulated more I_R, and have hence more $_sI_R$ to extinguish. This is not impossible, but there is of course no direct evidence to support such a long chain of arguments.

A consideration of PRD might be of assistance. Frith (1969) has suggested a possible cause for PRD in his work on strategies in skilled motor performance. Working with simple tapping tasks, Frith found that some subjects choose a strategy which results in maximum performance, leading to the accumulation of I_R and the occurrence of involuntary rest pauses (IRP's). Others prefer to work at submaximal levels, avoiding the build-up of I_R and the occurrence of IRP's; in consequence they may also, according to theory, avoid the growth of $_sI_R$ to any significant degree. If we can assume that subjects high in ability (whether due to innate or environmental factors) may be motivated to try harder for maximum performance from the beginning, they might accumulate sufficient I_R to lead to performance decrement even during pre-rest practice; having accumulated I_R and having experienced a series or IRP's they would also accumulate $_sI_R$ which could then be extinguished post-rest, giving rise to PRU as in Eysenck's theory. PRD would follow as a consequence of continued maximum performance strategy. Low-ability subjects, on this account, would choose a different strategy, i.e. one of submaximum effort, thus avoiding I_R build-up, the occurrence of IRP's, and the genesis of $_sI_R$. Hence for this group there would be no PRU or PRD. If these strategies are characteristic of high- and low-ability subjects respectively, then they would also characterize their practice on the component tasks which Reynolds & Adams have suggested would lead to ability differences at the beginning of the pursuit-rotor learning; hence a differential build-up of $_sI_R$ becomes a possibility and equally its extinction during the first minute or two of practice, giving rise to the upswing phenomenon pre-rest. This account is of course highly speculative, but it would account for the phenomena observed.

It might also be hypothesized that the higher-ability groups, by and large, would be more highly motivated; motivation and success are unlikely to be entirely separated. Hence in groups not specifically motivated for pursuit-rotor performance reminiscence, which is known to be a function of motivation (Eysenck, 1964*a*), would be greater in high-ability subjects; this differential would be wiped out when an external motivating factor is introduced, bringing the low-drive subjects up to the level of the most highly (internally) motivated subjects. This additional hypothesis would explain the observed relationship between ability level and reminiscence in low-drive groups, and its failure to be observed in high-drive groups. Again the hypothesis appeals to characteristics in task and subject which are difficult to observe and measure, and which certainly have not so far been measured; the main point in

proposing such a highly speculative theory is of course that it may lead to further investigations geared more specifically than the present one to testing the various parts of the theory in question.

One final hypothesis to be considered is concerned with the vexed problem of measurement of reminiscence and performance. Bahrick *et al.* (1957) have shown how the size of target may affect scores, and how the use of any particular size of target may produce measurement artifacts. In this connexion the work of Humphries (1961) is very relevant; he used a rotor in which circular target areas, insulated from each other, surrounded a very small central disc; recording from each of these target areas was separate, thus enabling scores to be obtained simultaneously from target areas differing in size. In his fig. 5 (p. 217) he has plotted results for 5 min. of massed practice, and 2 min. of practice following upon a 5 min. rest pause. It is clear that the larger target sizes give results similar to those obtained by our high-ability subjects, while the smaller targets give results similar to those obtained by our low-ability subjects. The former show pre-rest upswing, pre-rest downswing, and marked reminiscence, while the latter show neither pre-rest upswing nor downswing, and no reminiscence – although the scores are derived from the same subjects on the same occasion! This similarity is striking, but it does not prove that the phenomenon is an artifact when observed in high- and low-ability subjects respectively. Our second experiment has shown that shifting from a small target to a large one does not obliterate or change the observed phenomena; they are more dramatic in the case of the larger target, but they are identical with those produced by the smaller target. Neither can it be said that much larger or much smaller targets would eliminate our findings; it would seem impossible to enlarge the target much beyond the larger one used by us, as with such a target the subjects would be on target almost 100 per cent of the time. The target could of course be reduced, but in that case very little in the way of scoring would be possible – even with our small target scores were very low. Our two targets seem to span a large part of the area likely to give any reasonable learning data, and hence we feel that the similarity of results may serve to reassure readers worried by the arguments of Bahrick *et al.* or the data furnished by Humphries. Future work will undoubtedly benefit from incorporating multiple targets into the design, as in the Humphries experiment; we doubt if such innovations will alter the main conclusions drawn from our results.

The senior author is indebted to the Maudsley and Bethlem Royal Research Fund for partial support of this inquiry.

REFERENCES

ADAMS, J. A. (1961). The second facet of forgetting: a review of warm-up decrement. *Psychol. Bull.* **58**, 257–273.

ALPER, T. G. (1948). Task-orientation and ego-orientation as factors in reminiscence. *J. exp. Psychol.* **38**, 224–238.

AMMONS, R. B. (1947). Acquisition of motor skill. I. Quantitative analysis and theoretical formulation. *Psychol. Rev.* **54**, 263–281.

BAHRICK, H. P., FITTS, P. M. & BRIGGS, G. E. (1957). Learning curves: facts or artifacts. *Psychol. Bull.* **54**, 256–268.

BUXTON, C. E. (1943). Level of mastery and reminiscence in pursuit learning. *J. exp. Psychol.* **32**, 176–180.

Buxton, C. E. & Grant, D. A. (1939). Retroaction and gains in motor learning. II. Sex differences and a further analysis of gains. *J. exp. Psychol.* **25**, 198–208.

Ciectat, V. J. & Noble, C. E. (1958). Ability vs. practice in two hand coordination. *Percept. mot. Skills* **8**, 226.

Clark, A. M. (1967). The relationship of arousal to reminiscence, attention, withdrawal and the Hunter Prognostic index in chronic schizophrenia. (Unpublished Ph.D. thesis, University of London.)

Eysenck, H. J. (1956). 'Warm-up' in pursuit rotor learning as a function of the extinction of conditioned inhibition. *Acta psychol.* **12**, 349–370.

Eysenck, H. J. (1962). Reminiscence, drive and personality: revision and extension of a theory. *Br. J. soc. clin. Psychol.* **1**, 127–140.

Eysenck, H. J. (ed.) (1964*a*). *Experiments in Motivation*. Oxford: Pergamon Press.

Eysenck, H. J. (1964*b*). An experimental test of the 'inhibition' and 'consolidation' theories of reminiscence. *Life Sci.* **3**, 175–188.

Eysenck, H. J. (1965). A three-factor theory of reminiscence. *Br. J. Psychol.* **56**, 163–181.

Farley, F. H. (1966). Reminiscence and post-rest performance as a function of length of rest, drive and personality. (Unpublished Ph.D. thesis, University of London.)

Feldman, M. P. (1963). A reconsideration of the extinction hypothesis of warm-up in motor behaviour. *Psychol. Bull.* **60**, 452–459.

Fleishman, E. A. (1960). Abilities at different stages of practice in rotary pursuit performance. *J. exp. Psychol.* **60**, 162–171.

Frith, C. D. (1969). Individual differences in pursuit rotor and tapping skills. (Unpublished Ph.D. thesis, University of London.)

Gray, J. E. (1968). Level of arousal and length of rest as determinants of pursuit rotor performance. (Unpublished Ph.D. thesis, University of London.)

Humphries, M. (1961). A method of estimating 'RMS error' in pursuit rotor performance. *Percept. mot. Skills* **13**, 211–225.

Jahnke, J. C. (1961). Postrest motor learning performance as a function of degree of learning. *J. exp. Psychol.* **62**, 605–611.

Jones, M. B. (1966). Individual differences. In E. A. Bilodeau (ed.), *Acquisition of Skill*. London: Academic Press.

Leavitt, H. J. (1945). The relation of speed of learning to amount retained and to reminiscence. *J. exp. Psychol.* **35**, 134–140.

Leavitt, H. J. & Schlosberg, H. (1944). The retention of verbal and of motor skills. *J. exp. Psychol.* **34**, 404–417.

Locke, E. A. (1965). Interaction of ability and motivation in performance. *Percept. mot. Skills* **21**, 719–725.

McNemar, Q. (1933). Twin resemblances in motor skills, and the effect of practice thereon. *J. genet. Psychol.* **42**, 70–97.

Reynolds, B. & Adams, J. A. (1954). Psychomotor performance as a function of initial level of ability. *Am. J. Psychol.* **67**, 268–277.

Vandenberg, S. G. (1962). The hereditary abilities study: hereditary components in a psychological test battery. *Am. J. hum. Genet.* **14**, 220–237.

Zeaman, D. & Kaufman, H. (1955). Individual differences and theory in a motor learning task. *Psychol. Monogr.* **69**, no. 6. (Whole no. 391.)

(*Manuscript received* 17 *December* 1969; *revised manuscript received* 23 *April* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 217–223
Printed in Great Britain

HAPTIC AESTHETIC VALUE OF THE GOLDEN SECTION

By J. M. HINTZ and T. M. NELSON
Department of Psychology, University of Alberta, Edmonton, Alberta, Canada

The study re-examined the finding that blind and sighted subjects haptically perceiving various rectangular proportions, including the golden section, preferred the square as most pleasing. While rectangular preferences of congenitally blind subjects questioned the existence of the golden section as a haptically satisfying figure, the preferences of late blind and sighted subjects generally confirmed the haptic aesthetic value of the golden section. The results thus indicated that the haptic perception of the golden rectangle as aesthetically pleasing is contingent on contact with the visual world.

With the exception of Révész (1950), empirical studies investigating the aesthetic value of the golden rectangle have assumed that the simple figure is a visually satisfying form (see Woodworth, 1938; Valentine, 1962). Révész pursued the question of whether the golden rectangle was open to haptic aesthetic appreciation. The Danish psychologist simultaneously presented Fechner's (1897) series of 10 rectangles (25 sq.cm.) as well as a series of 26 larger and 26 smaller rectangles to blind and sighted subjects perceiving purely haptically. He found that, with few exceptions, subjects selected not the golden section, but the square, or a rectangle very similar to the square. Révész concluded that preferences for squares was not connected with an aesthetic impression, but was determined by the fact that the square occupies a position of its own as metrically the simplest rectangular figure. He also noted that the square is the easiest rectangle for the hand to grasp and apprehend by movements following the outline, that it has the simplest kinaesthetic rhythms, and that it is most easily retained in memory.

From further extensive research on haptic perception and the aesthetic experience of haptics with blind and sighted subjects, Révész concluded that in the field of haptics, aesthetics is limited to the most simple structures and part-structures, since the necessary and sufficient condition underlying aesthetic experience was the unified and spontaneous apprehension of form. He thus reduced the aesthetics of haptics to sensory impressions perceived in the stimulus structure of simple forms, and ignored the possible influence of the perceiver's visual organization as a determinant of *any* aesthetic judgement based on haptic perception.

The blind psychologist Cutsforth (1933) pursued another tack. Although he maintained that there are probably conditions in a given tactual stimulus pattern that should produce common affective values for sighted and blind alike, he held that 'the amount of visual organization present determines the subjective nature of every tactual experience' (p. 183). The subjective nature of any tactual experience for a congenitally blind person who lacks visual imagery would therefore differ from that of a blind person who lost his sight after having contact with the visual world.

The study to be reported was designed to study form preferences of both blind and sighted subjects in order to subject the phenomenon to further investigation. Specific hypotheses were not proposed. However, it should be noted that the type of analysis Révész provides predicts that (*a*) sighted subjects perceiving haptically and visually

will aesthetically prefer most frequently the square and the golden section respectively as the most pleasing rectangular figures, and that (*b*) congenitally and late blind subjects perceiving haptically will aesthetically prefer most frequently the square as the most pleasing rectangular figure. Reasoning from Cutsforth's position leads to a prediction that differences among congenitally blind, late blind and sighted subjects will occur, with the latter two groups being similar to, but different from, the congenitally blind.

METHOD

Subjects. Four groups of subjects served. The groups were: (1) congenitally blind (n = 20: nine males, 11 females); (2) late blind (n = 20: 11 males, nine females); (3) blindfolded normal sighted (n = 40: 20 males, 20 females); and (3*a*) non-blindfolded normal sighted (n = 20: 10 males, 10 females). The non-blindfolded sighted subjects were randomly selected from the blindfolded sighted subjects.

The following criteria were used in selecting blind subjects from the Canadian National Institute for the Blind: (*a*) a chronological age of 17 years or higher; (*b*) ophthalmologically diagnosed as possessing no more than light perception, and (*c*) absence of evidence of brain damage. Onset of blindness for congenitally blind subjects was at birth or before 5 years of age (Lowenfeld, 1963, p. 231). Onset of blindness for late blind subjects was after 7 years of age. In actuality, all congenitally blind subjects were blind before the age of 2½ years. All late blind subjects were not blind prior to 10 years of age.

Normal sighted subjects making up groups 3 and 3*a* were drawn from University of Alberta students and University of Alberta non-academic employees so as to be matched for age and sex with the blind subjects of groups 1 and 2.

Apparatus. Visual and tactual rectangular outlines with constant areas of 81 sq.cm. but having width–length ratios of 0·10, 0·20, 0·30, 0·40, 0·45, 0·50, 0·55, 0·60, 0·65, 0·70, 0·75, 0·80, 0·90 and 1·00 served as stimuli.* Actual size of the figures ranged from 2·84 × 28·46 cm to the square having sides of 9 cm. The specific value of 81 sq.cm. was chosen in order both to provide an adequately large area for haptic exploration and to make contact with prior research. Révész used a series of 'large' rectangles in his study. We do not know the actual size of the stimuli, however, since he did not report the precise dimensions of his rectangles. Visual rectangles were drawn with black ink on 25·4 × 30·48 cm off-white paper. Tactual rectangles were engraved in 25·4 × 30·48 cm durable gloss paper, their outlines were filled with contact cement, and they were mounted on cardboard. The raised figures were 1 mm wide, and the surface figure texture as well as the ground was smooth.

The overriding desire in design was to select stimuli and conditions adequate from the standpoint of all groups involved while offering special advantages in terms of overall breadth. This will inevitably restrict research to some of the details that may be important to specific groups. Thus both tactual and visual rectangles were presented in the frontal parallel plane with the long side horizontal, using a 30 cm square box which secured targets on four sides. This orientation was chosen so that investigation of sighted preferences might be made complete with visual field measurements as suggested by Stone & Collins (1965). These results are reported elsewhere, however (Hintz & Nelson, 1970).

Visual rectangles were inspected by group 3*a* at a distance of 101·6 cm and behind a 60·96 × 63·5 cm black cardboard field-stop. Peepholes 12 mm in diameter were cut into this so as to allow binocular viewing with eyes fixed at a height representing the centre point of the target. The target was illuminated by fluorescent fixtures providing 23 ft-c. on the target surface.

Procedure. The subjects were comfortably seated at a table in front of the swivel box. They were instructed that the experiment was an investigation into preferences for rectangles which would vary in proportion. Then they were told they would be presented with three cards, each containing an outlined rectangle and should select the rectangle most aesthetically pleasing, next most pleasing, and least pleasing. Instructions specified that this was to be done by 'touching and feeling the outlined rectangles with their preferred hand'. These directions lead most

* In a pilot study it was determined that congenitally blind, late blind and blindfolded sighted subjects could tactually discriminate a rectangular width–length ratio difference of 0·05 cm. Thus the smallest width–length ratio difference between two rectangles was 0·05 cm.

subjects to use both the fingers and the whole hand. The aesthetic component was stressed when answering questions. No time limits were imposed in exploration. These general instructions were used for the visual task following a 5-min. period for photic adaptation. Tactual preferences for subjects in group 3*a* were determined first to make the primary tactual data comparable.

Aesthetic preferences were determined by a process of successive approximation. Subjects were initially presented with rectangles having ratios of 0·30, 0·60, 0·80 and first, second and third choices recorded. The criterion for preferences in the first set was the rectangle chosen two-thirds of the time in a minimum of six trials. Presentation positions were randomized. Following this, a new set of rectangles representing proportions falling in the range of the first and second choices were offered and the same choice procedure used. Using this technique, it was possible to determine each subject's preference within a range of five rectangles by trial block 12. These five rectangles were then presented three at a time in all possible combinations for 10 trials. The rectangle preferred most frequently in the final trial block was taken to represent the subject's preference.

The rule of thumb followed continuously throughout this procedure was the successive elimination of those ratios not in the immediate surround of the subject's first and second preference, while maintaining a range of possible ratios to choose from which maximized the importance of the direction of the second preference. However, when the subject's first and second preferences were extreme with respect to one another, for example 0·30(1), 0·80(2) and 0·60(3), or 0·80(1), 0·30(2) and 0·60(3), the succeeding range of ratios presented to the subject centred on his first preference. In this example the second set of rectangles presented to the subject would have been 0·10, 0·40 and 0·55, or 0·65, 0·75 and 1·00 respectively.

Results

Aesthetic preferences of all subjects were found to cover the entire range of width–length ratios from 0·10 to 1·00. No sex differences were evident. Table 1 shows the modal and median values for all groups. Notice that late blind tactual preferences are closest to golden section values. Sighted subjects are very nearly the same for tactual as for visual forms, but both somewhat less than the late blind. The congenitally blind group is still more deviant in this regard.

Table 1. *Mode and median values for haptic and visual preferences*

	Congenitally blind	Late blind	Blindfolded sighted	Non-blindfolded sighted
Mode	0·10	0·60	0·60	0·60
Median	0·500	0·615	0·575	0·558

Fig. 1 presents results graphically in the form of frequency polygons. Note that in the limits 0·55–0·65, which includes the golden section, the percentage of non-blindfolded sighted, blindfolded sighted, late blind and congenitally blind subjects preferring these ratios is 40, 37·5, 30 and 10 per cent respectively. The polygon of congenitally blind subjects thus shows a clustering of preferences at the 'narrow form' end.

From the findings just cited, it appears that sighted subjects perceiving haptically as well as visually preferred an approximate golden section rectangle over the square. The Wicoxon matched-pairs signed ranks test (Siegel, 1956, pp. 75–81) was used with these data (see Table 2) to determine whether haptic and visual preferences of the sighted group differed significantly in this respect. Analysis was performed with each subject acting as his own control. Since the calculated value of $T = 27{\cdot}5$ was not

equal to or less than 21, the critical value required for a two-tailed test of significance at $\alpha = 0{\cdot}05$, the null hypothesis would not be rejected. This analysis is thus inconsistent with the results of Révész, who found that sighted subjects perceiving rectangles haptically and visually preferred as most pleasing the square and the golden section respectively.

Because congenitally blind, late blind and blindfolded sighted subjects did not most frequently prefer the square, it was decided to test for significant deviations from the golden section. Accordingly, the Kruskal–Wallis one-way analysis of

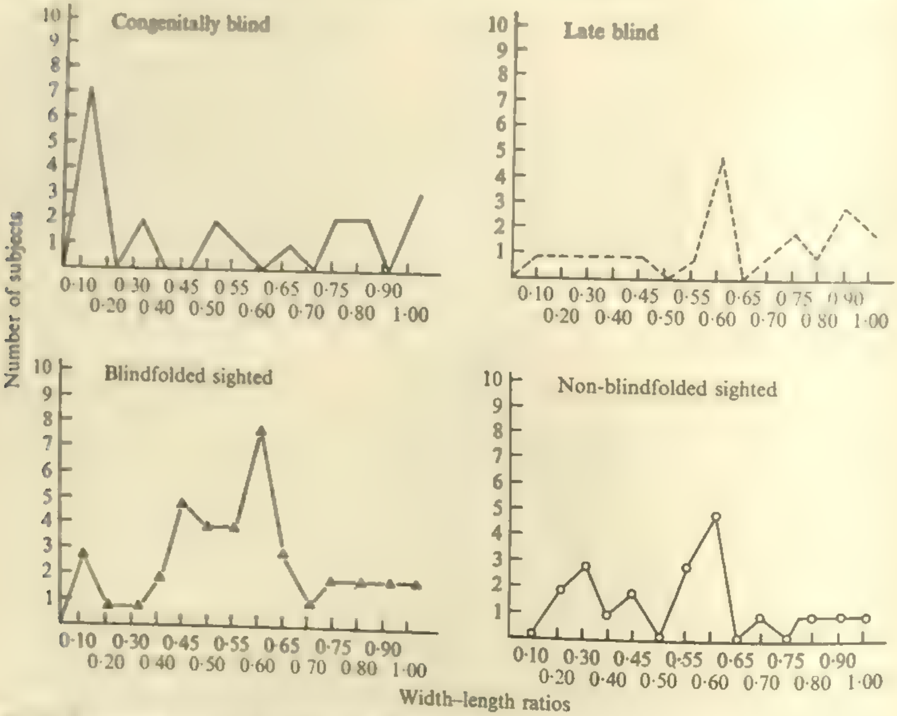


Fig. 1. Frequency polygons comparing preferences of congenitally blind, late blind, blindfolded sighted, and non-blindfolded sighted subjects for the most aesthetically pleasing rectangle.

Table 2. *Haptic and visual rectangular preferences of 20 normal sighted subjects*

Subject	Preference ratio: Haptic	Preference ratio: Visual	Subject	Preference ratio: Haptic	Preference ratio: Visual
1	0·40	0·45	11	0·40	0·40
2	0·55	0·45	12	0·50	0·55
3	0·45	0·30	13	0·60	0·60
4	0·70	0·70	14	0·50	0·30
5	0·65	0·60	15	0·90	0·80
6	1·00	0·90	16	0·60	0·60
7	0·45	0·20	17	0·45	0·30
8	0·60	1·00	18	0·65	0·60
9	0·20	0·20	19	0·60	0·55
10	0·45	0·55	20	0·60	0·60

variance by ranks (Siegel, 1956, pp. 185–94) was applied to the data. Results disclosed that with $H = 8{\cdot}733$, d.f. $= 2$ and $\alpha = 0{\cdot}05$, $P < 0{\cdot}02$. Thus the proportion preferences of congenitally blind, late blind and blindfolded sighted subjects differ reliably from one another in their deviations from the proportions of the golden section, and the null hypothesis would be rejected. This analysis is also inconsistent with the position taken by Révész but is consistent with the prediction of Cutsforth, so far as these can be developed.

Discussion

Preference data obtained from late blind and blindfolded sighted subjects generally confirmed the haptic aesthetic value of the golden section. However, preferences of congenitally blind subjects call into question the existence of the golden section as a haptically satisfying simple figure.

With the partial exception of late blind subjects, the preferential data cast doubt upon the conclusions of Révész (1950). Although late blind subjects preferred the golden section (see Fig. 1), 25 per cent also preferred rectangular width–length ratios of 0·90 or 1·00. All other subjects operating on haptic rather than visual input did most frequently prefer the simplest rectangular figure, or some rectangle very similar to the square. It thus seems highly questionable whether the 'metric principle' plays the decisive role that Révész attributes to it in determining preferences for simple rectangular forms perceived haptically.

Alternatively, it could be considered that the non-equivalent findings of the present study and those of Révész were due to a difference between the stimuli and the method used in the two studies. It is not possible to compare in detail the findings of the two studies, however, since Révész seems only to have reported his results very generally in *Psychology and Art of the Blind* (1950, p. 199).

The fact that preferences of congenitally blind, late blind and blindfolded sighted subjects significantly differed from one another in their deviations from the golden section is attributable to the fact that 35 per cent of the congenitally blind subjects preferred a rectangular width–length ratio of 0·10.

Visual imagery may be one factor determining these differences. Löwenfeld (1963) cites several studies (Toth, 1930; Schlaegel, 1953; Blank, 1958) showing that individuals who have lost their sight before the age of between 5 and 7 years do not retain any useful visual imagery. Related to this, Sylvester (1913) used a form-board test on 85 blind subjects and found that those lacking visual experience showed the least ability, and that the longer each subject had been able to retain his vision, the more successful he was in the test. The latter declares that 'those who have had visual experience retain their visual imagery and are assisted by it in the interpretation of their tactual impressions; and tactual imagery, even in those who have no other resource, is not as effective as a combination of tactual and visual imagery, (p. 200).

Likewise, Worchel (1951), in attempting to determine the role of visualization in tactual form perception, found that even though there were no significant differences between congenitally blind, late blind and sighted subjects in their ability to recognize simple geometric forms by selecting among four blocks the one that was similar in shape to the stimulus block (a finding confirmed by Ewart & Carp, 1963), groups did

differ in their ability to reproduce their tactual perceptions with drawings as well as describe them. Sighted subjects were significantly better than the congenitally blind. He also interprets this as indicating that touch alone is not as efficient in the perception of simple tactual forms as touch aided by visual images.

In this study it was generally observed that while late blind subjects made substantial use of visual imagery in reporting reasons for their rectangular preferences, congenitally blind subjects did not. The following examples of verbal reports illustrate this observation. Congenitally blind: 'Fatter shapes represent durability, quality, stability.' 'Just like the way it is, it's nice.' 'Don't really know; I like it.' 'Less rigid, a free flowing form, like a voice that flows.' 'A square has no beauty.' Late blind: 'I visualized picture frames and books; it's the size of a snapshot.' 'The proportions are pleasing, like electrical switch covers.' 'There is more use for a narrow rectangle, like a nameplate or a door slot.' 'Perceived it with reference to a frame; it's like a playing card.' 'Like linoleum tiles—they all fit; you get uniformity—nothing is offset.'

In noting the biological importance of vision as a special sense, Critchley (1953) writes that 'even in those who lose their sight at a relatively early age, a visual type of thinking may continue or at any rate, it may modify the kind of imagery' (p. 27).

In light of the work and conclusions of others, it seems consistent to interpret our findings as evidence that choice of the golden section as an aesthetically pleasing haptic rectangular figure probably depends upon experience with the visual world sometime in life. Haptic perception *per se* does not find any special value inherent in this particular rectangular proportion.

Additional support for this conclusion is offered by the developmental research of Thompson (1946) and Shipley *et al.* (1947), who studied preferences for rectangular proportions. Both these studies emphasize the important role experience with the visual world has in influencing the development of aesthetic preferences for rectangular proportions.

Finally, it would be remiss not to mention that the association of the golden section with art works has been periodic and certainly not universal, although the section has certainly appeared in the history of art more often than can be accounted for by mere coincidence (Scott, 1951, p. 64; Praeger, 1958, p. 260; Gardner, 1959, pp. 489, 421; Huyghe, 1959, pp. 127, 84; Bergamini, 1963, pp. 96, 97; Barry *et al.*, 1964, pp. 304, 47; Arnheim, 1966, pp. 106–10). If we should assume that the golden rectangle when present in art occurs chiefly as a consequence of sighted preferences, then perhaps the numerous empirical studies investigating the visual aesthetics of the golden section have been measuring a sighted population stereotype even when studying 'aesthetic' judgements of enlightened artists and designers. Great caution is needed here, however, because approximate golden section ratios do occur very frequently in nature (Scott, 1951, pp. 56, 59).

It still remains to offer a tentative explanation of why congenitally blind subjects (35 per cent) most frequently preferred a rectangular width–length ratio of 0·10.

Cutsforth (1933), adventitiously blind himself, has keenly observed that the blind frequently produce self-stimulation kinaesthetically through 'bodily swaying, rolling or tilting the head, arm motion and shoulder shrugs, and exaggerated genuflections' (p. 7). This automatic self-stimulation he calls 'blindisms'. One might

suggest that the frequent preferences congenitally blind subjects show for long narrow rectangles may not have been based on the 'aesthetic' value of the structure *per se*, but upon an increase in the amount of kinaesthetic stimulation received through tracing the outline of the figure. In short, the haptic perceptual exploration of long narrow rectangles provides more greatly contrasting kinaesthetic rhythms than the simple kinaesthetic rhythms of the square. Cutsforth (1933) expresses much the same idea when he says, 'there is much evidence from preliminary investigation that changing form possesses the greatest tactual aesthetic value' (p. 184).

References

Arnheim, R. (1966). *Art and Visual Perception: a Psychology of the Creative Eye.* Berkeley: University of California Press.

Barry, Sir Gerald, Bronowski, J., Fisher, J. & Huxley, Sir Julian (1964). *Man the Artist: His Creative Imagination.* London: MacDonald.

Bergamini, D. & the Editors of Life. (1963). Picture essay: the mathematics of beauty in nature and art. *Life Science Library: Mathematics.* New York: Time Inc.

Blank, H. R. (1958). Dreams of the blind. *Psychoanal. Q.* **27**, 158–174.

Critchley, M. (1953). Tactile thought, with special references to the blind. *Brain* **76**, 19–35.

Cutsforth, T. D. (1933). *The Blind in School and Society: a Psychological Study.* New York: American Foundation for the Blind, 1951.

Ewart, A. G. & Carp, F. M. (1963). Recognition of tactual form by sighted and blind subjects. *Am. J. Psychol.* **76**, 488–491.

Fechner, G. T. (1897). *Vorschule der Aesthetik*, vol. 1. Leipzig: Breitkopf & Härtel.

Gardner, H. (1959). *Art through the Ages.* New York: Harcourt, Brace.

Hintz, J. M. & Nelson, T. M. (1970). Golden section: reassessment of the perimetric hypothesis. *Am. J. Psychol.* **83**, 126–129.

Huyghe, R. (1959). *Ideas and Images in World Art.* New York: Harry N. Abrams.

Löwenfeld, B. (1963). Psychological problems of children with impaired vision. In W. M. Cruickshank (ed.), *Psychology of Exceptional Children and Youth*, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.

Praeger, F. A. (1958). *The Praeger Picture Encyclopedia of Art.* New York: Author.

Révész, G. (1950). *Psychology and Art of the Blind.* London: Longmans, Green.

Schlaegel, T. F., Jr. (1953). The dominant method of imagery in the blind as compared to sighted adolescents. *J. genet. Psychol.* **83**, 265–277.

Scott, R. G. (1951). *Design Fundamentals.* New York: McGraw-Hill.

Shipley, W. C., Dattman, P. E. & Steele, B. A. (1947). The influence of size on preferences for rectangular proportion in children and adults. *J. exp. Psychol.* **37**, 333–336.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Stone, L. A. & Collins, L. G. (1965). The golden section ıevisited: a perimetric explanation. *Am. J. Psychol.* **78**, 503–506.

Sylvester, R. H. (1913). The mental imagery of the blind. *Psychol. Bull.* **10**, 210–211.

Thompson, G. G. (1946). The effect of chronological age on aesthetic preferences for rectangles of different proportions. *J. exp. Psychol.* **36**, 50–58.

Toth, Z. (1930). *Die Verstaltungswelt der Blinden.* Leipzig: Barth. (Cited in Löwenfeld, 1963.)

Valentine, C. W. (1962). *The Experimental Psychology of Beauty.* London: Camelot Press.

Woodworth, R. S. (1938). *Experimental Psychology.* New York: Holt.

Worchel, P. (1951). Space perception and orientation in the blind. *Psychol. Monogr.* **65**, no. 15, 1–27.

(*Manuscript received* 29 *January* 1970; *revised manuscript received* 28 *April* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 225-233
Printed in Great Britain

A STUDY OF IMPERFECTLY ACQUIRED VOCABULARY

By KATE LOEWENTHAL

Department of Social Science and Humanities, City University, London

Three experiments examined behaviour with words which could be used and defined by the subjects, words which were simply 'familiar', and unknown words. Known and familiar words were recognized equally readily after brief tachistoscopic presentations; known words elicit semantically related responses, while familiar and unknown words usually elicit phonologically related or phonologically mediated responses. In a word-selection task, the subjects had to choose the word most similar to the stimulus, from a set of four words: semantically related, phonologically related, phonologically mediated and unrelated. For familiar words a semantically related word was more likely to be recognized in this choice task than it was to be produced in the word-association task. The same was true but to a lesser extent for unknown words. A simple scheme for word-processing is proposed to fit the findings, in which there are separate stores for the surface-structure (phonological/visual) features and dictionary entries of words. For known words there is ready access to entries in the surface-structure stores, and ready access between surface-structure and dictionary-entry stores. Access to any entries which may exist is very limited in the case of unknown words. For familiar words, there is ready access to surface-structure entries, but access to the dictionary-entry store is limited.

The experiments in this paper were not done with any strong hypothetical framework in mind. They were explorations of the fact that people seem to understand a great many more words than they actually produce in their own discourse. However, the experiments are not concerned with the distinction between words which a person has and has not used, but with the distinction between words which people feel they *could* use, and with words which they do not feel they could use.

More specifically, three stages in vocabulary acquisition are proposed. Initially a word is completely unfamiliar. Next, it is familiar: the subject could not define it but perhaps could understand it in an appropriate linguistic context. Finally, a word is known: it could be defined and it could be used appropriately. Words in the second of the three stages of acquisition constitute what is called the recognition vocabulary, which comprises words which might be understood in a context of connected discourse, but which cannot be produced readily. Words in the third stage constitute the active or expressive vocabulary.

The experiments all used intelligent adult subjects, and investigated the differences in treatment between members of these three classes of words in tachistoscopic recognition, in a word-association experiment, and in a situation in which the subjects have to select from several words one which is most similar to the stimulus.

Experiment I

Tachistoscopic recognition

This experiment is a variant of the 'familiarity and recognition' study first carried out by Howes & Solomon (1951). Twenty low-frequency words were presented singly for tachistoscopic recognition, and were subsequently classified by each subject into one of three classes. The difference between the present experiment and the traditional familiarity and recognition experiment stems from the fact that the traditional

experiments rest their interpretation on the assumption that frequently occurring words will be more familiar to the subject. This is no doubt likely, but the present experiment uses the more direct method of getting individual subjects' ratings of familiarity for words which are equated for frequency of occurrence.

Method

Subjects. These were 12 undergraduates (five males, seven females) in the University College of North Wales, Bangor. (The results from two further men had to be discarded on the ground that, for both of them, none of the words fell into the unknown class.)

Materials. Twenty seven-letter nouns, occurring at least once per 4 million, but not as often as once per million, were taken from the Thorndike Lorge (1944) count. The words were: lamprey, acolyte, bustard, primacy, spicule, benison, nostrum, filbert, transom, widgeon, vulgate, lampoon, sextant, atavism, bitumen, culvert, doeskin, rowlock, verbena, probity. Preliminary testing indicated that for the university educated subject, fairly equal numbers of words fall into each of the three classes. The words were typed in lower-case lettering in the centre of 3 × 5 in. cards for tachistoscopic presentation.

Procedure. The subject was seated in front of the tachistoscope and it was explained that a word would be flashed repeatedly on the screen until he was able to recognize it. The subject was warned that some words might be unfamiliar and he was encouraged to guess at the word no matter how silly it sounded. The exposure was always 0·1 sec., and two practice words (servant, honesty) were presented to eliminate warm-up effects. Subsequently the 20 words were presented for recognition, in a different randomized order for each subject. If no recognition occurred after 15 exposures, the subject was told that the exposures of that word would be discontinued, and he was told what the word was. After the 20 words had been exposed, the subject was given a list of the words with the following instructions: 'Could you place an A, B or C by each of the words listed below, according to the following scheme: A, you know what the word means, could define it, and could use it in the correct way. B, the word is familiar, but you are not sure of its meaning. C, you have never heard of the word, and have no idea what it means.' Preliminary testing indicated that subjects found this classification scheme made sense, and they found little difficulty in applying it. Finally, the subject was asked for comments and introspective reports on the experiment.

Results

The quantitative analysis was done by taking the median number of exposures required for recognition, for the words in each class for each subject. (Means could not be used because of occasional failures to recognize the word after 15 exposures.) A Friedman two-way analysis of variance on the medians showed a significant effect of knowledge of words ($\chi_r^2 = 141{\cdot}8$; d.f. $= 2$; $P < 0{\cdot}001$), although examination of the means of the medians, which are given here for convenience of exposition, suggests that the difference lies between known and familiar words on the one hand, and unknown words on the other. The means were: known (Kn) words: 2·7 exposures; familiar (F): 3·1 exposures; unknown (U) words: 4·5 exposures.

Discussion

There is little difference between the number of exposures required for successful recognition of known and familiar words. Even though subjects say they cannot use or define the F words, they seem to have acquired them to the extent that they can match information from brief presentation of the words with their vocabulary store as successfully as for words which cannot be defined. This feat of matching is less frequently accomplished with unknown words, although sometimes subjects are able to recognize quite quickly words which they say are unknown. This may be because

all the necessary information has been got from the tachistoscopic exposures, or it may be that subjects have some knowledge of the word without being aware of it. The subjects seemed to spend a few seconds after each exposure in a memory search, and or in word construction, before announcing a guess or their inability to make a guess. Some comments were: 'I got the general look of the word and then searched my memory to find something like it.' 'Between the exposure and the reply, you've got a vague idea and are shuffling through to find a word. You have, say, three consonants and are trying to find vowels to fit.' Sometimes the information is correctly matched with stored vocabulary; sometimes it is incorrectly matched (e.g. EPICURE for spicule). Sometimes a guess is constructed which is both an approximation to English and an approximation to the word presented (e.g. ATAVIAN, STAVIAN, STEVISM, STAVISM guessed successively by one subject, who finally succeeded in 'recognizing' atavism). A final resort is to decipher the word letter by letter.

The finding of the experiment confirms the usefulness of calling the set of familiar words the recognition vocabulary: these words can be recognized as quickly as words in the expressive vocabulary.

Experiment II

Word-association

Children's word-associations typically differ from those of adults. There is a greater frequency of 'clang' or phonetic associations, and also a greater frequency of syntagmatic responses. These latter are responses of a different grammatical class from that of the stimulus, and it is said of them that they can be accounted for by having occurred in contiguity with the stimulus (Ervin, 1961); examples would be good–BOY, deep–HOLE, as opposed to good–BAD, deep–WIDE, which are paradigmatic responses typical of adults.

The present experiment was done with the idea of comparing the sorts of word-associations given to known, familiar and unknown words. It was thought that childlike types of association might occur to imperfectly acquired vocabulary, but no specific hypotheses were made as to the exact stage of acquisition at which they would disappear.

Method

Subjects. The subjects were 20 undergraduates (11 males, nine females) in the University College of North Wales, Bangor. None of them had served as subjects in Expt. I.

Materials. These were the same 20 low-frequency words used in Expt. I.

Procedure. Instructions for the word-association task were to 'respond to each word with the first word that occurs to you', and the subject was warned that some words might be unfamiliar. Three warm-up words were then presented aurally (servant, monocle, diamond) followed by the 20 low-frequency words. A different randomized order of presentation of these 20 words was given to each subject. After the word-association task, the subjects were asked to classify the words, as described in Expt. I. Finally, the subjects were questioned about any bizarre associations, and asked to 'give some indication, if you can, of why you gave that response'.

Results

On 12 occasions no associations were given. The remaining 388 responses were classified as S (semantically related), P (phonologically related), PM (phonologically mediated) and X (unrelated).

Semantically related responses (S) (n = 146) were related to the publicly acceptable 'meaning' of the stimulus, e.g. acolyte–PRIEST, MASS, INCENSE. There were a few responses (n = 17) which were both semantically and phonologically related to the stimulus, e.g. acolyte–PROSELYTE, bustard–BUZZARD. One suspects that these responses might not have occurred were it not for the double semantic and phonological relationship; these responses were classed as semantically related.

Phonologically related responses (P) (n = 92), e.g. acolyte–ACCOLADE, ACORN; benison–BENZENE, VENISON.

Phonologically mediated responses (PM) (n = 106) are superficially bizarre, but they can be accounted for by interpolating a mediating word between stimulus and response; the mediator is phonologically related to the stimulus and semantically related to the response, e.g. lampoon–(harpoon)–WHALING, culvert–(cul-de-sac) -END. The mediating words were not guesses on the part of the author; they were readily supplied by the subjects when they were being questioned about their responses, at the end of the experimental session.

Unclassifiable responses (X) (n = 44) could not be included in any of the above categories, e.g. culvert–APPLE.

Table 1. *Frequency of different types of word-association, classified subjects' knowledge of stimulus words*

Knowledge of stimulus	Semantically related	Phonologically related	Phonologically mediated	Unrelated	Total
Known	111 (82·2%)	8 (5·9%)	14 (10·4%)	2 (1·5%)	135 (100%)
Familiar	29 (24·0%)	29 (24·0%)	43 (35·5%)	20 (16·5%)	121 (100%)
Unknown	6 (4·5%)	55 (41·7%)	49 (37·1%)	22 (16·7%)	132 (100%)
Total	146	92	106	44	388

Table 2. *Component χ^2 calculated from the data of Expt. II*

	Value of χ^2 (using theoretical parameters)	D.F.	Significance level
χ^2_k (knowledge of stimuli)	4·51	2	n.s.
χ^2_r (response type)	8·14	1	$P < 0·005$
χ^2_s (subjects)	5·09	19	n.s.
χ^2_{kr}	161·08	2	$P < 0·001$
χ^2_{ks}	85·92	36	$P < 0·001$
χ^2_{rs}	56·03	19	$P < 0·001$
χ^2_{krs}	48·32	36	n.s.
Total χ^2	369·09	115	

Table 1 shows the frequencies of these different types of responses, classified according to subjects' estimations of their knowledge of the stimulus words. The frequencies are pooled over all subjects for presentation here, but for the purposes of statistical computation, the data were cast in three-dimensional matrix (knowledge of stimulus × response type × subjects). Because cell frequencies were small, X responses were omitted from the analysis, and P and PM response frequencies were pooled, giving a 3 × 2 × 20 matrix. Values of χ^2 were computed using Lancaster's (1951) method (described by Lewis, 1962), using theoretical parameters. Table 2 summarizes the results of the analysis: interaction between stimulus knowledge and

response type gives the largest component χ^2; highly significant values of χ^2 were also found for responses, stimulus knowledge × subjects, and responses × subjects.

Discussion

The main finding of the experiment is that Kn words are typically responded to with semantically related associations, while typical responses to F and U words are phonological or phonologically mediated. F words are slightly more likely than U words to elicit semantically related responses, but this tendency is so very slight that in this word-production situation F and U words are virtually indistinguishable. F words are inadequately connected with a semantic system as far as word-production goes.

In the light of the analogy with children's word-associations it had been expected that some syntagmatic responses would occur. However, there was an absence of such responses. There were a very few responses which were not nouns and which were classed as semantically related responses, but it seems dubious to call them syntagmatic responses, since it seemed unlikely that they could be accounted for by having occurred in contiguity with the stimulus. Rather, they seemed to be attempts at producing a synonym, which had failed to be nominalized (e.g. atavism–GENETIC, lampoon–DECRY). This lack of syntagmatic responses agrees with McNeill's (1966) account of word-associations. He suggests that paradigmatic responses cannot be accounted for by extending the 'experience of words in contiguity' explanation, as suggested by Ervin (1961), but that they 'arise through the same process that creates productive distribution classes'. Word-associations are words which have the maximum number of features in common with the stimulus. Some of these features are semantic, and some are syntactic. Many apparently syntagmatic responses of children occur because children do not know many of a word's features; some of the missing features may be syntactic. Hence the occurrence of 'pseudo-syntagmatic' responses, which are in fact genuine attempts to generate a word with the maximum number of features in common with the stimulus, e.g. eat–FOOD. In the present experiment the lack of syntagmatic responses can be accounted for by the adult subjects having adequate knowledge of the syntactic features of the stimulus; they are thus able to match syntactic features of stimulus and response successfully.

Experiment III

Selection of a similar word

In the word-association task familiar words are unlikely to elicit semantically related responses. This experiment tests the hypothesis that when the subject is confronted with a word which is semantically related to a 'familiar' stimulus he is likely to recognize the similarity between the words, although a semantically related response is not likely to be *produced* in word-association.

Method

The subjects were confronted with a choice of four types of word, one for each of the four types of association given in Expt. II: semantically related, phonologically related, phonologically mediated and unrelated. The subject's task was to select the word most similar to the stimulus.

Subjects. These were 36 undergraduates (30 male, six female) in the City University, London. Three subjects were run individually, the rest being run in four group sessions.

Materials. The same 20 nouns were used as stimuli as in the previous experiments. The words from which selection was to be made were, wherever possible, words which had actually been given as responses in Expt. II (e.g. probity–INTEGRITY, PROGRESS, LAWYER, ELASTICITY). The stimulus was typed in upper-case lettering, and alongside it in lower-case lettering were the four words whose order was determined using random number tables.

Procedure. The subjects were given two duplicated sheets containing all the instructions and experimental words. The first sheet instructed the subject to classify the 20 low-frequency words into three classes, as in the earlier experiments. This was done before the experimental task, as it was likely that the subject's knowledge of the words might be affected by carrying out the Expt. I task.

The second sheet was headed with the instructions: 'On the left-hand side of the page are the words you have just seen, in block capitals. Please underline the *one* of the four words which you think is *most closely related* to the word in capitals on the left. Only underline one word from each set of four. Obviously the task is difficult because you will probably be unfamiliar with many of the words in the left-hand column (although they are all real words)—but please make a guess for each word. Ask if you have any questions.' Afterwards, for all except one of the groups, the subjects' comments were invited, and the purpose of the experiment was discussed.

Results

The quantitative results are shown in Table 3, which is similar in layout to Table 1. Component χ^2 were calculated as in Expt. I, and Table 4 summarizes the results. Highly significant values of χ^2 are obtained for the effects of knowledge of stimulus, response type, stimulus knowledge × response type, and knowledge of stimulus × subjects.

Table 3. *Frequency of different types of word-choice, classified by subjects' knowledge of stimulus words*

Knowledge of stimulus	Semantically related	Phonologically related	Phonologically mediated	Unrelated	Total
Known	193 (93·7 %)	2 (1·0 %)	11 (5·3 %)	0	206 (100 %)
Familiar	92 (60·1 %)	15 (9·8 %)	28 (18·3 %)	18 (11·8 %)	153 (100 %)
Unknown	172 (50·8 %)	52 (15·4 %)	47 (13·9 %)	67 (19·8 %)	338 (100 %)
Total	457	69	86	85	697

Table 4. *Component χ^2 calculated from the data of Expt. III*

	Value of χ^2 (using theoretical parameters)	D.F.	Significance level
χ^2_k	45·36	2	$P < 0·001$
χ^2_r	149·03	1	$P < 0·001$
χ^2_s	14·82	35	n.s.
χ^2_{kr}	47·69	2	$P < 0·001$
χ^2_{ks}	138·53	70	$P < 0·001$
χ^2_{rs}	46·27	35	n.s.
χ^2_{krs}	86·30	70	n.s.
Total χ^2	528·00	215	

Discussion

The hypothesis of the experiment is confirmed in that when subjects classify a word as merely familiar, they are far more likely to select a semantically related response as similar than they are to produce a semantically related association. The subjects

commented: 'If it was a word where I wasn't sure of the meaning, it jumped out as soon as I saw it on the second page' and 'When you're doing the second sheet it suddenly dawns on you what the words are'.

The surprising thing is that subjects are almost as likely to select semantically related responses to U words as to F words. Part of the explanation may lie in the subjects' guessing strategies. It appeared from their comments that the subjects wanted to avoid being tricked into selecting phonological and phonologically mediated responses, although they were not always successful in detecting the latter. Thus many guesses were in effect between two words rather than four. Even so it would be expected that guesses should be evenly divided between the semantically related and unrelated responses. That this is not so suggests that even when people claim a word is unfamiliar they may have sufficient knowledge about it for them to be able to recognize a word which is unrelated in meaning. Thus, having eliminated 'nostril' and 'pulpit' (mediated by 'rostrum') as suitable choices for the stimulus 'nostrum', the subject decides that 'nostrum' is unlikely to be the name of a fish, and so eliminates 'herring', leaving him with 'remedy'.

A SCHEME FOR WORD-PROCESSING

The data from the experiments could be accounted for in several ways. Here it will be suggested that information about words is stored in a system with separate components for semantic and surface-structure (phonological/visual) information. Fig. 1 sets out the components of this system as it might exist in a literate monolingual.

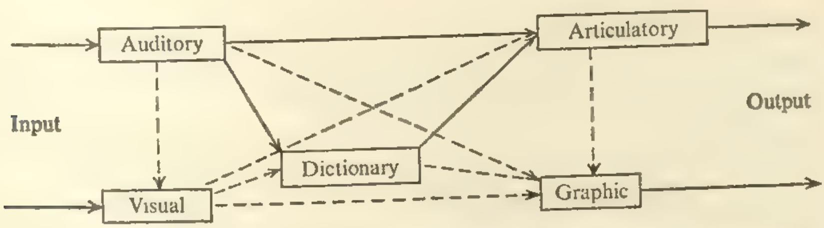


Fig. 1. Components of a word-processing scheme. (Dubious connexions are indicated by dotted lines.)

The notion of separate input and output components is perhaps uneconomical but it is one way of accounting for selective impairment of expressive and receptive abilities in aphasia. There are also separate components for different encoding and decoding skills when these are well-practised: there may be specific alexia or agraphia, leaving speech comprehension or production unimpaired, and *vice versa*. The arrangement of surface-structure information is unlikely to be random. It may be that entries for lexical items are arranged in a network with entries for similar items adjacent. More efficient is the proposal that there is an auditory dictionary entry for each word, with features such as stress pattern, number of syllables, phonemic clusters arranged in order of specificity (see, e.g., Brown & McNeill, 1966). Entries on the input side need not be complete, but must be sufficient to give access to information in the dictionary and the output stores about the same lexical item. Entries in the output surface-structure stores must be complete for articulation to be possible.

Semantic information could be stored in a dictionary of the sort proposed by Katz & Fodor (1963), in which entries consist of syntactic and semantic markers, distinguishers and selection restrictions. McNeill (1966), Brown & McNeill (1966) and Marshall & Newcombe (1966) have applied this theory to performance in word-association, the tip-of-the-tongue state, and dyslexia.

The findings of the experiments can be accounted for by arguing that in the case of imperfectly acquired vocabulary—familiar words—entries exist in the surface-structure stores and the dictionary, but access between stores is limited. Known words have entries in both kinds of store, and there is ready access between entries, while unknown words have incomplete or non-existent entries in any store.

Known words are recognized readily after tachistoscopic exposure. Information is matched to an entry in the visual store, which is matched to an entry in the articulatory store. There is no evidence from this experiment as to whether or not the dictionary is by-passed, i.e. whether understanding is an accompaniment of recognition. (More detailed models of word-recognition performance are provided by Sperling (1967) and Morton (1969). These authors explore many complexities overlooked by the present discussion.) In word-association and word choice, the auditory/visual entry gives ready access to the dictionary component, resulting in the production or selection of a word with a similar dictionary entry.

Familiar words are recognized as readily as are known words from tachistoscopic exposure, and probably much the same processes are involved in both cases, although it seems more probable that the dictionary entry is by-passed in the recognition of familiar words. In word-association the subjects try to produce a response with the maximum number of features in common with the stimulus, with semantic matching taking priority over phonological matching. Because access to the dictionary entry is limited in the case of familiar words, the common features are phonological. However, the semantic matching tendency seems to be strong, and so once phonological matching is achieved, the dictionary is entered and a phonologically mediated response is given. In the word-choice experiment semantic matching is achieved with familiar words, because the semantically related choice gives direction to the dictionary search. The 'lost' dictionary entry for the familiar word is more likely to be scanned than it would otherwise be: the meaning of the stimulus 'jumps out' at the subject, and he is able to perform semantic matching.

Performance with unknown words in recognition is poor presumably because of the lack of stored information with which to match the input: the stimulus has to be reconstructed by the subject. In word-association, phonological and phonologically mediated matching occurs in much the same way as with familiar words. The successful 'recognition' by subjects of the semantically related choice for unknown words in Expt. III may be partly accounted for by the elimination strategy previously described, together with the suggestion that subjects may have some fund of information which can be applied to words which they claim are unknown to them.

To recapitulate: stages in vocabulary acquisition may be related to the way information about words is entered in the components of a system with separate surface-structure and dictionary-entry stores. Known words have entries in the surface-structure and dictionary-entry components, and there is ready access from the entry in one store to the entry in the other store. Thus a word can be understood and

defined, and, given a definition, could be produced. Words which are familiar but which cannot be defined apparently have entries in the surface-structure components, and, according to Expt. III, some sort of dictionary entry. However, according to Expt. II, there is no ready access from the surface-structure to the dictionary entry. Probably (as Brown & McNeill have suggested) there is also poor access from the dictionary entry to the surface-structure entry: given a definition, there is trouble in producing a word. There is some evidence that there may be entries of some sort, even if very incomplete, for some of the words which people claim are unknown to them. Access to these entries, if and when they exist, is very limited.

Grateful acknowledgements are due to the Medical Research Council for the award of a Junior Research Fellowship held at the University College of North Wales, Bangor, where part of this research was carried out. Thanks are due to Steve Miller of the City University for statistical advice, and to Dr P. N. Johnson-Laird and other members of the Psychology Department, University College London for their interest and helpful comments.

References

Brown, R. & McNeill, D. (1966). The 'tip-of-the-tongue' phenomenon. *J. verb. Learn. verb. Behav.* **5**, 325–337.

Ervin, S. M. (1961). Changes with age in the verbal determinants of word association. *Am. J. Psychol.* **74**, 361–372.

Howes, D. H. & Solomon, R. L. (1951). Visual duration threshold as a function of word probability. *J. exp. Psychol.* **41**, 401–410.

Katz, J. J. & Fodor, J. (1963). The structure of a semantic theory. *Language* **39**, 170–210.

Lancaster, H. O. (1951). Complex contingency tables treated by the partition of chi-square. *J. R. statist. Soc.* B **13**, 242–249.

Lewis, B. (1962). On the analysis of interaction in multidimensional contingency tables. *J. R. statist. Soc.* A, **125**, 88–117.

McNeill, D. (1966). A study of word association. *J. verb. Learn. verb. Behav.* **5**, 548–557.

Marshall, J. C. & Newcombe, F. (1966). Syntactic and semantic errors in paralexia. *Neuropsychol.* **4**, 169–176.

Morton, J. (1969). Interaction of information in word recognition. *Psychol. Rev.* **76**, 165–178.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Sperling, G. (1967). Successive approximations to a model for short-term memory. *Acta psychol.* **27**, 285–292.

Thorndike, E. L. & Lorge, I. (1944). *The Teacher's Word Book of* 30,000 *Words.* New York: Columbia University.

(*Manuscript received* 21 *February* 1969; *revised manuscript received* 9 *March* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 235–243
Printed in Great Britain

THE ROLE OF IMAGERY IN INCIDENTAL LEARNING

By PETER W. SHEEHAN

Department of Psychology, University of New England, Armidale, N.S.W., Australia

Incidental (INC) and intentional (INT) subjects were given one of two orientating tasks requiring response to 18 high-imagery (I) and 18 low-imagery (N-I) noun stimuli equated for *m* and frequency of usage. The subjects either responded with a set to image stimuli at inspection (IMG set) or responded to stimuli in terms of their familiarity (FAM set). Independent groups of 18 subjects were allocated to each of the four experimental conditions and groups were equated for sex of the subject; the subjects were tested for free recall and recognition of I and N-I stimuli. It was predicted that if imagery plays a special role in INC learning, the difference in recall of I and N-I words would be greater for INC as compared to INT learning. The hypothesis was supported for FAM set in recognition where I and N-I stimuli were most differentiated. Type of orientating task was related to the presence or absence of the predicted effect.

In their discussion of subject-characteristics as related to success in incidental learning, Plenderleith & Postman (1956) emphasized that the subject maintains attention to multiple aspects of the stimulus, and has available differential responses to the stimulus items which vary in their effectiveness. Response habits with which the incidental learner enters the experimental situation appear to be the chief determinants governing the selectivity of incidental learning (Postman *et al.*, 1955). Ability to arouse sensory imagery is one such habit the subject brings to the task of incidental learning which has received no systematic analysis to date.

Sheehan & Neisser (1969) unexpectedly found that visual images accompanying recall of incidental designs were much more vivid than those accompanying designs which the subjects knew beforehand they would later be asked to remember. Vividness of imagery used in the recall of the incidental designs was also related significantly to the accuracy of their reproduction. There were several possible explanations for the difference in results with the two kinds of material: tasks for incidental and intentional recall were different; number of presentations varied, as did also the time between presentation and test of retention. Although the limitations of the design prevent one from drawing the inference that incidental learning procedures were responsible for the pattern of findings, results were suggestive of a positive association between incidental learning and vividness of imagery. The notion that imagery may play a special role in incidental learning has been supported most directly by Ernest & Paivio (1969). Employing a Type II design (Postman, 1964) they found that high-imagery subjects were superior to low-imagery subjects in their imagery for the irrelevant components of colour-word units.

Considerable work on the role of imagery in verbal recall has been reported by Paivio and his associates (Paivio, 1966; Paivio *et al.*, 1968; Yuille & Paivio, 1967). Their evidence suggests that verbal symbolic processes and images are equally available as mediators when the stimulus to be learned is concrete; but only verbal mediators are readily aroused when stimuli are abstract (Yuille & Paivio, 1967). The prediction that ease of learning depends on the image-arousing properties of individual nouns has been confirmed in studies using free recall (Dukes & Bastian, 1966;

Paivio, 1967) and recognition (Gorman, 1961) as measures of learning. In these studies learning was better for concrete or high-imagery arousing nouns than for abstract or low-imagery arousing nouns.

The differential recall of high-imagery (I) and low-imagery (N-I) nouns in recall would allow a test of the special role of the mediating capacity of imagery in incidental learning. Differential effects are known to be associated with abstract and concrete stimuli; concrete nouns are the more effective stimuli for arousing images which function as mediators in response recall (Paivio, 1969). Incidental (INC) learning can be distinguished from intentional (INT) learning in the availability to the subject of concrete imagery as a strategy of recall. In INC learning representational responses such as verbal coding are limited in function (Postman, 1964) but sensory imagery is one coding system which is readily available to the subject when he is tested for recall. Mediation of recall by imagery rather than by verbal coding is more likely to occur where representational responses are restricted as in INC learning than where verbal and imaginal mediators are both freely available strategies as under INT instruction. Accordingly, if imagery plays a particular function in INC learning it follows that the pattern of I/N-I response after subjects have received 'incidental' instruction will differ from the pattern of response the subjects will give after receiving 'intentional' instruction. It is predicted that there will be a greater difference between the number of I and N-I stimuli recalled in INC learning than in INT learning. The higher difference score for INC learning would indicate that I stimuli as distinct from N-I stimuli arouse imagery which plays a more specific function in INC learning than it does in INT recall.

A Type I design was chosen for the study, since interest was predominantly in associative processes as determined by the nature of stimulus materials, and this design is uncomplicated by response competition and generalization effects (Postman, 1964). Imagery is viewed in the experimental context as one aspect of the cognitive abilities which the subject brings to the testing situation and which functions more easily for some stimuli than for others. Imagery may be expected to be available to influence recall especially where the subject has not intentionally employed strategies of memory to code material at inspection. The design compares INC learning with INT learning for subjects who were given one of two orientating tasks: subjects either associated imagery to each stimulus as it appeared at inspection (Imagery set), or rated stimuli in terms of how frequently they were used in everyday speech (Familiarity set).

The following hypothesis was formulated: More I than N-I stimuli will be recalled in INC learning as compared to INT learning. Orientating task was varied to investigate whether or not this effect will occur independently of the kind of instructions given to subjects at inspection.

METHOD

Subjects. Four groups of 18 subjects were tested in a 2×2 factorial design. Independent sets of subjects were randomly allocated to one of the following experimental conditions, with the provision that each group was equated for the sex of the subject: (i) INT learning instruction without set for imagery (Familiarity set); (ii) INT learning instruction with set for imagery; (iii) INC learning instruction without set for imagery; and (iv) INC learning instruction with set for imagery. The subjects were drawn from a volunteer sample of first-year psychology under-

graduate students. Each experimental group contained nine males and nine females. The subjects were recruited by a notice which requested them to take part in an experiment concerned with the measurement of individual differences in rating scores on a number of different variables.

Stimuli. Eighteen high-imagery (I) and 18 low-imagery (N-I) nouns were chosen. Stimuli were selected for Imagery value and equated for meaningfulness (*m*) from the word list of 925 nouns rated for I (on a seven-point scale) and *m* by Paivio *et al.* (1968) and are listed in Table 1. The mean I values of the high- and low-imagery lists were 6·25 and 3·08, respectively. Mean *m* values for the two lists were 6·08 and 6·07, respectively. Stimuli were also equated for word frequency (*f*). The average *f* of words on the high- and low-imagery lists occurring in four and a half million instances were 192·2 and 195·1, respectively. Frequencies were taken from the word lists compiled by Thorndike & Lorge (1944).

Table 1. *Experimental stimuli* ($n = 36$)

High-imagery words (I)		Low-imagery words (N-I)	
acrobat	monk	attitude	knowledge
cigar	nail	chance	magnitude
dollar	physician	development	moral
engine	pianist	economy	necessity
factory	pipe	evidence	perjury
flag	tomb	explanation	position
harp	slipper	history	theory
newspaper	snake	immunity	welfare
microscope	tweezers	justice	gravity

Procedure. Stimuli were printed in matt-black ink on 5 × 3 in. plain index cards and exposed, tachistoscopically, for 4 sec. The subjects were given 16 sec. to rate either the vividness of their imagery, or their familiarity with each stimulus. The subjects under the imagery set (IMG) were asked explicitly to evoke an image of what the stimulus noun represented, to press a key when they had made their judgement and then to indicate the vividness of their imagery by writing down an appropriate rating selected from a seven-point scale for measuring vividness of visual imagery (Betts, 1909). The subjects under the familiarity set (FAM) were given comparable instructions and indicated their ratings on a seven-point scale of familiarity situated in front of them. Ratings for each stimulus were recorded by the subjects on separate pages of 40-page score booklets. The latencies of the subjects' ratings were recorded by a chronoscope which measured the period between stimulus onset and rating judgement. The timer was automatically started when the stimulus flashed on and was stopped by the subject pressing a key to indicate that he had made the judgement required. The stimulus was displayed for the full 4-sec. period, irrespective of whether or not the subject had pressed the key by that time. After pressing the key, the subject wrote down his rating in the score booklet; the experimenter then said 'ready?' and presented the next stimulus for inspection.

The experimenter instructed all subjects initially that this was a study on individual differences in rating scores. Instructions for INC and INT subjects under their respective sets (IMG and FAM) were identical, except that INT subjects were told additionally before testing began that: 'when I have shown you the last word and you have rated the vividness of your imagery (the frequency of your use) I am going to ask you to recall as many of the words as you can. This is also, then, a learning task.' All subjects were presented with two practice stimuli, 'tablespoon' (I) and 'sentiment' (N-I), and familiarized with the rating procedure before testing proper.

The experimental list was randomly arranged with respect to I and N-I stimuli, with the restriction that the last four and first four nouns alternated I and N-I words. The subjects were given one of two lists: random arrangement of I and N-I stimuli, or random arrangement reversed. The subjects within each experimental group were alternated from one list to the other. Thirty sec. after the last item had been presented, the subjects were asked to recall as many words as they could remember in any order. The period of recall was 5 min. A post-experimental inquiry was then given to determine if any INC subjects had a set to learn. The subjects were asked the following questions: (i) 'Did you expect that I would ask you to recall these words?' (If *Yes*: 'When did you expect that I would be asking you to recall?') (ii) 'Did you attempt to learn any of these words?' (iii) 'Did any particular words get your special attention during the

rating task?' (iv) 'Did you say any of the words over to yourself at all?' Following the standard inquiry, the subject answered questions about those stimuli which were recalled correctly by him. For each stimulus remembered the subject was asked to rate the vividness of the sensory image (if relevant) aroused by the word at the time of recall. The subject rated his imagery on a seven-point vividness scale (Betts, 1909), irrespective of his orientating task at inspection.

After this inquiry had been completed a recognition test was carried out. The subjects were presented with a test list of 100 nouns which included the 36 experimental stimuli. The remaining 64 stimuli consisted of 32 I and 32 N-I stimuli randomly interspersed among the experimental stimuli. The subject's task was to encircle those nouns which he recognized as having seen before. Finally, a 35-item questionnaire was administered to each subject to assess his capacity to arouse mental imagery independently of his experimental performance on the noun list. The test examined imagery in seven different modalities. It is reliable (Sheehan, 1967*b*) and has been shown by psychometric analyses to measure a general capacity to evoke sensory imagery (Sheehan, 1967*a*).

RESULTS

Two INC subjects indicated in the post-experimental inquiry that they thought they would be asked to recall the words. Results for these subjects were excluded from analyses and two additional subjects tested to equate groups in size.

Table 2. *Mean reaction times for FAM and IMG ratings to I and N-I stimuli for subjects tested under INC and INT instruction*

Orientating task	I value	Instruction: INC	Instruction: INT
IMG	I	1·40	1·87
	N-I	2·62	2·92
FAM	I	3·16	3·36
	N-I	3·23	3·55

An analysis of reaction-time data for ratings given at inspection was conducted to determine whether the two orientating tasks and the two sets of instruction yielded comparable time scores. A triple classification, split-plot analysis (Type III design, Lindquist, 1953) was carried out to analyse latency scores for orientating task (FAM, IMG), instruction (INC, INT) and stimulus imagery value (I, N-I). Table 2 sets out the mean reaction times for judgements given by subjects in the separate experimental groups. Analysis indicated significantly faster reaction-time scores for the IMG conditions as compared with the FAM conditions (for main effect of set: $F = 13{\cdot}18$; d.f. = 1, 68; $P < 0{\cdot}001$), and significantly faster times for I as compared to N-I stimuli (for main effect of stimulus I value: $F = 4{\cdot}20$; d.f. = 1, 68; $P < 0{\cdot}05$). The INC conditions were comparable to the INT conditions in reaction time (for main effect of instruction: $F < 1$; d.f. = 1, 68; $P > 0{\cdot}10$). Separate split-plot analyses of variance of reaction time data for the FAM and the IMG orientating tasks indicated that I stimuli were associated with faster rating times than N-I stimuli for the IMG task ($F = 46{\cdot}46$; d.f. = 1, 34; $P < 0{\cdot}001$), but not for the FAM task ($F = 2{\cdot}38$; d.f. = 1, 34; $P > 0{\cdot}10$). These results indicate that INC and INT instructions yielded comparable presentation time scores and may be legitimately compared for recall effects, but that the predicted effects should be examined for the orientating tasks separately since IMG and FAM tasks differ in the time the subjects spent in judging I and N-I stimuli.

Table 3 sets out the number of I and N-I stimuli correctly learned under each set and type of instruction for tests of free recall and recognition. Data from this table indicate that I stimuli were differentiated from N-I stimuli more for some conditions than for others; differentiation was most consistent for the recognition test where significantly more I than N-I stimuli were learned under each of the four experimental conditions. Table 4 abstracts from Table 3 the relevant scores for test of the main hypothesis of the study; it lists the difference scores (number of I stimuli correctly recalled or recognized, less the number of N-I stimuli correctly recalled or recognized)

Table 3. *Mean number of I and N-I stimuli correctly recalled by each experimental group in tests of free recall and recognition*

Test	Orientating task	Instruction	Imagery value	
			I	N-I
Free recall	IMG	INC	5·61	5·56
		INT	6·89	6·61
	FAM	INC	6·94	5·00**
		INT	7·00	5·44*
Recognition	IMG	INC	17·22	14·33***
		INT	17·00	14·33***
	FAM	INC	16·50	13·00***
		INT	16·17	14·17***

* $P < 0{\cdot}05$, ** $P < 0{\cdot}01$ and *** $P < 0{\cdot}001$ for the difference between number of I and N-I stimuli recalled, or recognized for each of eight experimental comparisons.

Table 4. *Mean difference scores between INC and INT instruction for subjects tested in free recall and recognition under FAM and IMG sets*

Test	Orientating task	Instruction	
		INC	INT
Free recall	IMG	0·05	0·28
	FAM	1·94	1·56
Recognition	IMG	2·89	2·67
	FAM	3·50	2·00*

* $P < 0{\cdot}05$

for both sets of instruction. Table 4 shows that difference scores were greater for INC as compared to INT instruction for three of the four experimental comparisons. The pattern of I minus N-I response in INC learning did not differ significantly from the pattern of response for INT learning for the IMG set, but difference scores differed significantly for INT and INC groups in recognition under the FAM set (for recognition analysis: $t = 1{\cdot}74$; d.f. $= 34$; $P < 0{\cdot}05$). Considering results from both Tables 3 and 4, the hypothesis of the superiority of the differences between I and N-I stimuli for INC as compared to INT learning was supported for the test of learning (recognition) where I and N-I stimuli were differentiated most clearly.

The nature of the obtained predicted effect for FAM set in recognition may be clarified to some extent by examination of the individual patterns of the total number of I and N-I responses given by subjects under INC and INT instruction. It is

important to note that the obtained result was not a simple effect of instructional variation in the overall number of I or N-I stimuli recognized. For the FAM set comparisons between INC and INT instruction were carried out for the number of I and N-I stimuli correctly recognized (for I stimuli: $t = 0{\cdot}53$; d.f. $= 34$; $P > 0{\cdot}10$, for N-I stimuli: $t = 1{\cdot}07$; d.f. $= 34$; $P > 0{\cdot}10$). These results suggest that the obtained effect was not due simply to a significant increase in recognition of I stimuli under INC as compared to INT instruction. The trend of data in Table 3 suggests rather that I responses were maintained under the two sets of instruction, while N-I responses were fewer under INC than under INT instruction. Analysis, in addition, showed that it was the number of *correct* I stimuli recognized that was relevant to the effect. The four experimental conditions did not vary among themselves in the number of false I positives recognized ($\chi^2 = 3{\cdot}50$; d.f. $= 3$; $P > 0{\cdot}30$).

Questionnaire data provided the opportunity for investigation of individual differences among subjects in their capacity to evoke imagery. Over all subjects receiving INT instruction ($n = 36$), the number of correct responses given in free recall bore no relationship to subjects' scores on the visual imagery scale of the questionnaire ($r = 0{\cdot}10$), but there was a small but significant relationship for INC learning ($r = -0{\cdot}34$; $P < 0{\cdot}025$).* The imagery questionnaire failed to discriminate the subjects in their recognition performance.

Discussion

Results suggest that imagery plays a special role in INC learning. Its function, however, appears to be dependent both on the particular test of recall and the nature of the orientating task. The present hypothesis cannot be investigated where there is no difference between I and N-I stimuli since it is argued specifically that a difference which exists in one kind of learning (INC) will be greater than a difference which exists in another kind of learning (INT). If recognition yields that differentiation more than free recall, as the present study indicates, then recognition is the more sensitive test for the effects of imagery on INC learning.

It remains unclear whether or not the recognition process itself is important to the demonstration of the INC effect. There are some substantial differences between recognition and free recall which may be relevant. In recognition the subject is presented with the stimulus cue to invoke recall; this test may hence be interpreted as more susceptible to the subject's use of sensory imagery as a coding system since the stimulus cue would enhance the probability that I items will be correctly retrieved on the test trial. Where I stimuli were differentiated most, the effect of imagery on learning was observed. Enhancement of I stimuli in recognition, then, may indicate that this method of recall lends itself most efficiently to the subject's discovery and utilization of mediating imagery.

The present design employs the same subjects in tests of both recall and recognition. Consequently, recognition data may be affected by the subject's free recall as

* The negative correlation of the questionnaire scores with measures of recall performance expresses a positive relationship between variables. The sign of the correlation is an artifact of the direction of scoring on the imagery test. The subjects rated test items on a scale where 1 indicated 'perfectly clear' imagery and 7 indicated 'no imagery at all'.

well as by the experimenter's inquiry into the subject's recall performance. The sequence of conditions, however, was identical for INC and INT instruction. The effects of one test upon the subjects' performance in the other would be expected to be present for both sets of instruction. The possibility of differential effects for the two types of instruction remains, however, and awaits empirical analysis.

It has been well established (Postman, 1964; McLaughlin, 1965) that instructions to learn produce appropriate representational responses during stimulus presentation whereas under INC instruction this is less so. The nature of the representational responses aroused during presentation depends to a large extent on the kind of orientating task the experimenter uses to ensure that the subject attends to the stimuli, and orientating tasks may have a facilitative or interfering effect on INC learning (Saltzman, 1956). Just as the INC effects (as traditionally investigated) have varied with the orientating task employed, so it appears that the role imagery plays in INC learning depends on the way in which stimuli are presented to the subjects in the first instance. The association of imagery with stimuli at inspection is one task which fails to produce the predicted advantage of I over N-I stimuli in INC learning.

Under the FAM set where the predicted differences did occur one cannot explain results by appealing to variation in association value between I and N-I stimuli since both concrete and abstract nouns were equated for m and f. Verbal mediation processes as reflected by these variables do not account for the obtained differences in patterns of I and N-I response. Any verbal coding systems that were employed as representational responses by the subjects would have been functionally linked to I and N-I stimuli to an equivalent degree. Performance differences in recall of the concrete and abstract words are best explained, then, in terms of the varying capacity of the two types of stimuli to arouse sensory imagery.

It might be argued that the absence of effect for IMG set is explained by the hypothesis that, under explicit instruction to evoke imagery, the subject made effective use of imagery for the abstract N-I stimuli. Support for this line of thinking comes from the evidence of McDonald (reported by Paivio, 1969), who found that subjects asked to evoke imagery often stated that concrete imagery was aroused by one or both members of abstract pairs. However, such an explanation of the present results is untenable, since analysis of subjects' vividness of imagery ratings given to correctly recalled stimuli showed that there was a significant difference between the vividness of imagery associated with I and N-I stimuli for the IMG set (t = 17·47; d.f. = 34; P < 0·001). For each of the four experimental conditions the recall of I stimuli was associated with more vivid imagery than the recall of N-I stimuli. Results suggest that even when imagery instructions were given to the subjects, they still could not use imagery effectively to recall abstract stimuli.

Close analysis of reaction-time data suggests that comparison of the recall of I and N-I stimuli was invalidated under IMG set. Over both sets and types of stimuli, INC and INT instruction yielded comparable reaction times but separate analyses showed that reaction times associated with I and N-I stimuli were appreciably different under imagery instruction. It is possible that subjects may have required longer to acquire the same number of representational responses to the N-I stimuli, or that the abstract words were sufficiently longer than the concrete words to require from subjects substantially more scanning or processing time. It seems most

plausible, however, that under the imagery set the subjects had more time to employ representational responses to aid the recall of N-I as compared to I stimuli. This hypothesis is consistent with the higher recognition scores for I stimuli (see Table 3), since there is the definite possibility of decoding error for N-I stimuli where complex strategies have been used by subjects at inspection. This hypothesis is of special relevance when it is noted that five out of eight subjects in the IMG–INC condition who said in the inquiry that they repeated stimuli to themselves at inspection explicitly mentioned that they did so only for the abstract stimuli because they were difficult to image. No subject who said he repeated words in the FAM–INC condition reported that he did so for any particular type of stimulus. Whatever is the most appropriate interpretation of the time discrepancy under IMG set, it is clear that interpretation of the comparison of I and N-I stimuli in recall is considerably more unequivocal when the opportunity the subject has to employ representational responses at inspection is the same for both I and N-I stimuli irrespective of the nature of the representational response the subject actually brings to the task.

In conclusion, results suggest that under certain specifiable conditions imagery serves a particular adaptive function in INC learning. Under these conditions subjects who are limited in their use of strategies of learning rely especially on non-verbal, imaginal processes to mediate recall. The higher proportion of I to N-I responses given in recognition for the FAM set suggests that, for this particular orientating task, imagery serves an effective mediating function in INC learning. Further work should explicate this function and investigate whether it is the differentiation of I and N-I stimuli that is essential to the demonstration of the effect or whether the recognition process itself is important.

This study was supported in part by a research grant from the University of New England.

REFERENCES

BETTS, G. H. (1909). The distribution and functions of mental imagery. *Columb. Univ. Contrib. Educ.* **26**, 1–99.

DUKES, W. F. & BASTIAN, J. (1966). Recall of abstract and concrete words equated for meaningfulness. *J. verb. Learn. verb. Behav.* **5**, 455–458.

ERNEST, C. H. & PAIVIO, A. (1969). Imagery ability in paired-associate and incidental learning. *Psychon. Sci.* **15**, 181–182.

GORMAN, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *J. exp. Psychol.* **61**, 23–29.

LINDQUIST, E. F. (1953). *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mifflin.

MCLAUGHLIN, B. (1965). 'Intentional' and 'incidental' learning in human subjects: the role of instructions to learn and motivation. *Psychol. Bull.* **63**, 359–376.

PAIVIO, A. (1966). Latency of verbal associations and imagery to noun stimuli as a function of abstractness and generality. *Can. J. Psychol.* **20**, 378–387.

PAIVIO, A. (1967). Paired-associate learning and free recall of nouns as a function of concreteness, specificity, imagery, and meaningfulness. *Psychol. Rep.* **20**, 239–245.

PAIVIO, A. (1969). Mental imagery in associative learning and memory. *Psychol. Rev.* **76**, 241–263.

PAIVIO, A., YUILLE, J. C. & MADIGAN, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *J. exp. Psychol.* **76**, Monogr. Suppl. no. 1.

PLENDERLEITH, M. & POSTMAN, L. (1956). Discriminative and verbal habits in incidental learning. *Am. J. Psychol.* **69**, 236–243.

Postman, L. (1964). Short-term memory and incidental learning. In A. W. Melton (ed.). *Categories of Human Learning*. New York: Academic Press.

Postman, L., Adams, P. A. & Phillips, L. W. (1955). Incidental learning studies: II. The effects of association value and of the method of testing. *J. exp. Psychol.* **49**, 1–10.

Saltzman, I. J. (1956). Comparisons of incidental and intentional learning with different orienting tasks. *Am. J. Psychol.* **69**, 274–277.

Sheehan, P. W. (1967*a*). A shortened form of Betts' Questionnaire Upon Mental Imagery. *J. clin. Psychol.* **23**, 386–389.

Sheehan, P. W. (1967*b*). Reliability of a short test of imagery. *Percept. mot. Skills* **25**, 744.

Sheehan, P. W. & Neisser, U. (1969). Some variables affecting the vividness of imagery in recall. *Br. J. Psychol.* **60**, 71–80.

Thorndike, E. L. & Lorge, I. (1944). *The Teacher's Wordbook of* 30,000 *Words*. New York: Teachers College, Bureau of Publications.

Yuille, J. C. & Paivio, A. (1967). Latency of imaginal and verbal mediators as a function of stimulus and response concreteness-imagery. *J. exp. Psychol.* **75**, 540–544.

(*Manuscript received* 7 *November* 1969; *revised manuscript received* 22 *May* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 245–252
Printed in Great Britain

SITUATIONAL DETERMINANTS OF OPEN-FIELD BEHAVIOUR IN *MUS MUSCULUS*

By DAVID A. BLIZARD
The Jackson Laboratory, Bar Harbor, Maine, U.S.A.

A cross of eight inbred lines of albino mice was made in order to construct a heterogeneous sample representative of the laboratory species. An experiment incorporating the factors of light, noise, arena size, sex and days was performed using an open-field test standardized on the rat. The results obtained were compared with previous findings on inbred lines and F_1s. Ambulation was markedly affected by most independent variables, whilst defaecation seemed less sensitive in relation to them. The utility of multiple crosses of inbred lines is discussed.

Open-field behaviour has been extensively studied in domesticated forms of *Rattus norvegicus* (Broadhurst, 1960) and of *Mus musculus* (Wimer & Fuller, 1966; Candland, 1970). Most experiments in this area have been restricted either to a comparison of inbred strains and their hybrids or to the analysis of samples of animals drawn from stocks of an unknown and unrepeatable degree of genetic heterogeneity. For example, even if a particular animal supplier adheres to a system of mating designed to maintain genetic stability (e.g. Poiley, 1960), it is possible that the degree of heterogeneity of the population and the actual structure of the gene pool will vary systematically over a period of time depending on the size of the population.

If it is desired to generalize about a particular species, it seems a good experimental strategy to sample as widely from its gene pool as possible. If this is done, an important characteristic of an animal species, its genetic variability, is included in the calculation of the error term of a particular study. Wild animals are generally characterized by genetic heterogeneity, although its degree is affected by the relative frequency of small family groups with their limiting effect on genetic variability (Crowcroft, 1966). Thus the use of heterogeneous stocks of laboratory animals may tend to render results less idiosyncratic when they are compared with wild forms of the same species. A method of creating such heterogeneity in the laboratory is to breed multiple crosses of inbred strains (McClearn & Meredith, 1964). Because residual genetic variation in the inbred lines which contribute to the cross is close to zero, samples may be obtained on future occasions with the same average degree of heterogeneity. Such groups will also, with predictable sampling variation, have similar gene pools.

This experiment was concerned with open-field behaviour in a heterogeneous sample of laboratory mice. It has been suggested (Collins, 1966) that there are important differences in open-field behaviour between the rat and the mouse, but no studies have been performed which enable a comparison to be made between the behaviour of the two species in an identical stimulus situation. The experiment was therefore designed to investigate the behaviour of a multiple cross of inbred mouse strains in the open-field test standardized by Broadhurst (1960). Its specific aim was to evaluate the effects of variation in certain physical stimuli on the behaviour of a heterogeneous group of mice.

METHOD

Subjects. The eight inbred lines used as the base population for the experiment were mated in accordance with a scheme devised by Dr E. P. Lee. The albino strains A/HeJ, [illegible] BALB/cJ, 129/J, SJL/J, AKR/J, SWR/J and RF/J were obtained from The Jackson Laboratory [illegible]. Maternal effects were controlled by reciprocal mating only *after* the first set of crosses had been made as certain matings in the first cross were fertile in only one sire-dam combination. Litters of pups with a potential genetic contribution from all eight inbred lines were much larger than normal (mean, 10·2; range, 4–15), thus facilitating the generation of sufficient experimental animals. The genotypes of animals resulting from an eight-way cross will be composed on an average of one-eighth of the genotype of each inbred line, if a large enough sample of animals is taken. The precise degree of heterogeneity generated is impossible to assess as the inbred lines used are undoubtedly identical at many loci.

Apparatus. The open-field used has been described in detail elsewhere (Broadhurst, 1960). Briefly, it is a circular arena (diameter 32·75 in.) the walls (12·5 in. high) of which are painted flat white. Its floor is divided by concentric circles and radial lines into segments approximately equal in area. The arena is illuminated from above by light sources mounted on a frame which also serves as a support for a muslin screen. A loudspeaker is connected to the frame over the centre of the arena which enables a white noise stimulus to be delivered to the open-field.

Procedure. Animals were weaned at 28 days, then weighed and sexed. Thereafter they were kept with litter mates of the same sex (5–8 per cage) until tested between 81–90 days of age. Animals were maintained on a 12 hr. light–12 hr. dark cycle with a room temperature of approximately 74° F.

A factorial design was used to investigate the effects of variation in illumination level, sound intensity and arena size on open-field behaviour. Sex differences were also included for their intrinsic interest and behaviour was studied over four consecutive days, thus introducing a fifth factor (days) into the design. The factors of illumination (L) and noise (N) were at three levels, whilst sex (S) and arena size (A) were at two. Thus 36 animals were required for a single replication, the fifth factor, days (D), involving repeated measurements on the same subject. Three replications involving a total of 108 subjects were run independently, animals being assigned to treatments when they reached the appropriate age. The duration of the whole experiment was approximately one month.

Light and noise levels were selected on the basis of values used in a previous study on the rat (Broadhurst, 1957), although a third point was included for each variable to avoid the difficulty posed by estimating functional relationships based on two points. Illumination levels were thus 1, 30 and 165 ft-c. These were estimated on several parts of the floor of the open-field by a Gossen light-meter. White noise levels were 65, 78 and 94 db (with reference to 0·0002 dynes per cm²) at floor level as measured by a Scott Type 410 A sound level meter. The smaller arena used in the present study had no equivalent in the previous one (Broadhurst, 1957). It was one half the diameter of the larger arena. The variable was manipulated by placing a movable hardboard cylinder (diameter 16·375 in.) within the large arena on appropriate trials.

Two days before testing animals were marked on the tail with ink and moved to a room adjacent to the testing room. Testing was carried out in the latter half of the animal's light cycle, the order of testing being randomized within each day. Each animal was tested for 2 min. per day and a score for ambulation was obtained by tracing the path of the animal on a plan of the open-field. The total number of segments entered was noted afterwards. The defaecation score (the number of faecal boli deposited) was also recorded at the end of each trial. After each animal had been removed from the open-field, boli were removed and the floor wiped with a damp cloth.

RESULTS

The data were principally treated by the analysis of variance. Repeated measures on the same subject were taken into account by the use of a method described by Winer (1962). In the analyses for both ambulation and defaecation, the data were pooled after no significant differences were found between independent replications.

Tests used after the analysis of variance found significance were either Newman-Keuls or Duncan's multiple range test.

Ambulation. Part of the analysis relating to this category may be inspected in Table 1. In summary, significant main effects involving all variables except sex were demonstrated. In addition, significant AL and ALS interactions occurred. The group means relating to the main effects are contained in Table 2.

Table 1. *Analysis of variance of ambulation scores*

(Only significant F ratios ($P < 0.05$) are included for the sake of brevity.)

Source	S.S.	D.F.	M.S.	F	P
Arena (A)	56,948·2	1	56,948·2	83·459	<0·001
Illumination (L)	122,239·0	2	61,119·7	89·573	<0·001
Noise (N)	7,904·6	2	3,952·3	5·792	<0·01
Sex (S)	234·1	1	234·1	0·343	
AL	11,579·6	2	5,789·8	8·485	<0·001
ALS	9,829·6	2	4,914·8	7·203	<0·01
Error	49,129·3	72	682·4		
Days (D)	8,451·1	3	2,817·0	10·986	<0·001
Error	55,385·2	216	256·4		
Total	360,739·0	431			

Table 2. *Mean ambulation scores* in relation to main effects*

Arena size			Illumination level		
Large	50·4	$P < 0·001$	1 ft-c.	62·5 (1 v. 30, 165; $P < 0·001$: 30 v. 165; n.s.)	
Small	27·4		30 ft-c.	29·8	
			165 ft-c.	24·4	

White-noise levels		Days	
65 db	37·4 (94 v. 65, 78; $P < 0·01$: 65 v. 78; n.s.)	1	46·01 (day 1 v. 2, 3, 4; $P < 0·001$: no other sig. diff.)
78 db	34·6	2	39·32
94 db	44·7	3	35·63
		4	34·80

* Segments crossed per 2 min. trial.

The overall mean ambulation score was 38·9 entries. Animals ambulated more in the larger arena, ambulating almost twice the distance compared with the smaller arena. The single variable which appeared to have the greatest effect on ambulation was variation in light intensity. As light decreased, within the present limits, ambulation increased. Ambulation at 1 ft-c. was between two- and threefold the level at the highest light intensity. The illumination level of 1 ft-c. had a greater effect on ambulation than arena size or noise level as judged by the mean level (Table 2). Conversely, the highest illumination level produced the lowest mean ambulation score of all main effects.

In comparison, noise level had slight effects. This was somewhat surprising as the highest level of this variable approached the intensity used by Fuller & Collins (1968)

to elicit audiogenic seizures, which are partially characterized by wild running. No seizures were observed during the course of this experiment.

The decrease in activity over days was principally related to higher scores on day 1 compared with the following three days. The changes between days 2, 3 and 4 were in the same direction but became progressively smaller. The finding is generally compatible with previous work on inbred lines (Brookshire & Rieser, 1967).

Table 3. *Mean ambulation scores: interaction between arena size and illumination*

(E = expected mean value. O = observed mean value.)

	Illumination					
	1		2		3	
Arena	E	O	E	O	E	O
Large	56·45	81·28	40·1	36·9	37·4	33·0
Small	44·95	43·7	28·6	22·7	21·6	15·8

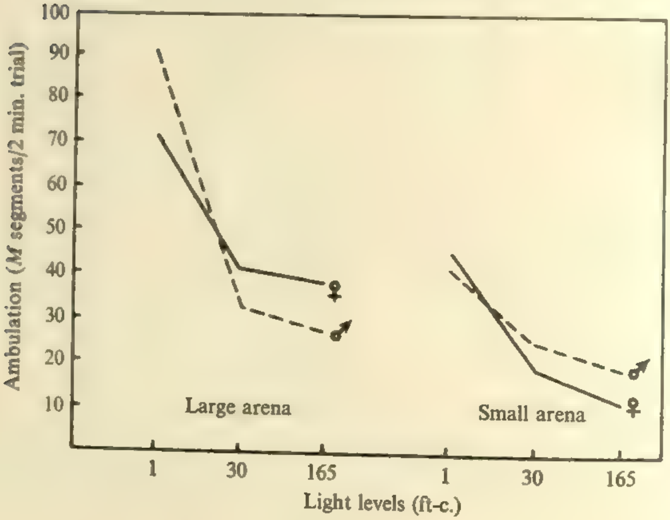


Fig. 1

The AL interaction is illustrated in Table 3. Mean values expected for the interaction assuming simple additivity of the main effects involved are included as well as the observed means. It appears that the large arena and lowest illumination level interacted markedly to increase ambulation scores (expected = 56·45; observed = 81·28). In all other combinations of these two variables the expected value is a slight overestimate of the obtained value.

In Fig. 1 the ALS interaction is presented graphically. In the small arena, the main effects appeared to combine in an additive manner if the AL interaction was taken into account in calculating the expected level of ambulation. In the large arena, however, at the lowest illumination level, the mean score of males differed markedly from expectations based on additivity of main effects, whereas females did not appear to diverge appreciably from expected levels. The sex difference observed at this

combination of levels tended to disappear or even reverse itself at higher light intensities in the large arena. However, neither sex diverged markedly from expectations based on additivity at these levels of illumination.

Defaecation. The experimental treatments were much less effective in influencing the defaecation response. Only illumination achieved significance as a main effect ($F = 3{\cdot}746$; d.f. = 2, 72; $P < 0{\cdot}05$). The means associated with this comparison were as follows: 1 ft-c., 2·02; 30 ft-c., 2·84; 165 ft-c., 2·76. Defaecation in 1 ft-c. of illumination was significantly less than in 30 and 165 ft-c. The two higher levels did not appear to affect defaecation scores differentially.

Table 4. *Sex × day interaction: mean number of faecal boli deposited per day*

	Days				
	1	2	3	4	
Males	2·52	1·89	2·09	2·59	(day 2 *v.* 1 and 4; $P < 0{\cdot}05$)
		< 0·01	< 0·01		
Females	2·39	2·94	2·91	3·0	

A significant SD interaction is described in Table 4 ($F = 3{\cdot}237$; d.f. = 3, 216; $P < 0{\cdot}05$). It appeared to have two sources. Females defaecated significantly more than males on days 2 and 3 ($P < 0{\cdot}01$) but did not differ significantly on days 1 and 4. There was, in addition, a significant decrease between days 1 and 2 for males and a significant increase between days 2 and 4. However, it should be noted that there was no significant difference between days 1 and 4 with regard to defaecation by males.

Discussion

Open-field behaviour

Ambulation. In general, the results on those factors in the design which have previously been extensively investigated with inbred mice agreed quite well with previous data. With regard to the effect of illumination level on activity, McReynolds *et al.* (1967) found that activity in the BALB/cJ strain decreased markedly from the condition of darkness to 4–5 ft-c., thereafter continuing to decrease but with much smaller decrements at higher illumination levels. This finding is in substantial agreement with the present study where the major decrement in ambulation occurred between the 1 ft-c. and 30 ft-c. levels, whilst there was no significant difference between the two higher levels of illumination. As far as the directional effects of illumination level on activity are concerned, the results on a multiple cross of albino strains appear similar to previous data established on inbred albino strains. In both cases, activity was reduced by a factor of one half as the illumination level changed from darkness or near-darkness to intermediate intensities (18–30 ft-c.). However, one limiting possibility concerning this comparison may be that the actual level of activity differs according to the nature of the genetic background. As the data presented by McReynolds *et al.* (1967) were presented in different measurement units and were transformed before means were calculated, no comparisons regarding possible level differences could be made.

The significant effect of arena size introduces certain problems of interpretation. The open-field test was suggested by Hall (1934) to represent a fearful stimulus to rodents. Thus it could be argued that larger arenas should provoke a more intense fear reaction than smaller arenas. However, ambulation in the large arena was twice the level observed in the small arena. If lower levels of ambulation are related to fear, then the reverse relationship should hold. However, several recent studies question the usual hypothesized negative correlation between fear and activity (Halliday, 1966; Blizard, 1968), so the result does not necessarily conflict with current theoretical formulations.

The effects of noise level on ambulation were less striking than the effects of variation in light intensity. Increase in noise beyond certain levels did have an effect on motor behaviour. It was noticeable, however, that the effect of noise seemed to be independent of the other situational stimuli manipulated. The results concerning noise may also be more capable of generalization to the laboratory mouse as a species. It would not appear that the albino locus is as important in relation to the reception of sound stimuli as it is known to be with regard to light.

The decrease in ambulation found over days parallels findings on inbred lines (Brookshire & Rieser, 1967). Again, the effect does not seem restricted to albino strains and so may represent a general characteristic of the laboratory species. The high ambulation observed in the mouse on the first day of testing may be partly the result of high levels of fear on initial exposure to the open-field. Whimbey & Denenberg (1967) found that, in the rat on the first day of testing, activity had a positive loading on an emotionality factor. From day 2 onwards it had a negative loading on this factor. If similar processes operate in the mouse, this could be an explanation of the decrease in ambulation over days.

The failure to find a significant main effect with regard to sex agrees with previous work on inbred strains (McClearn, 1960; Dixon & DeFries, 1968). The finding was tempered, however, by the significant ALS interaction (Fig. 1). At the lowest level of illumination in the large arena, males were significantly more active than females. That males are more sensitive to variation in situational stimuli was suggested by the tendency towards reversal of sex differences at the higher levels of light intensity in the large arena (Fig. 1). It is unknown whether these findings relate solely to sex differences in an eight-way cross. However, experimentation on inbred lines has incorporated similar levels of light intensity and arenas of similar size without the occurrence of such a sex difference.

Defaecation. The significant increase in defaecation with increasing light intensities found in the present study parallels previous work by Ross *et al*. (1966) and McReynolds *et al*. (1967) on inbred lines. The greatest change in defaecation was found when the 1 ft-c. level was compared with the two higher levels.

The lack of significance of an overall effect related to sex contrasts with previous findings on inbred lines (Henderson, 1967; McReynolds *et al*., 1967), where males were generally found to have higher scores than females. Henderson (1967) found that mice undisturbed before behavioural testing in adulthood exhibited greater sex differences in defaecation in the hybrid *v*. inbred form, a result previously obtained by Bruell (1964). This was interpreted as indicating that greater sex differences may be a population dominant characteristic. If such were the case, then one would have

expected sex differences in the same direction to have occurred in the present study where a heterogeneous sample was used. The failure to find sex differences in the same direction suggests that population dominance with regard to a particular character would be more reliably assessed if a larger variety of genetic interactions were sampled than occur in F_1 hybrids.

It has been suggested (Bruell, 1969) that murine defaecation may be correlated with territoriality. If this were the case, higher defaecation scores might be predicted in males and also increases in defaecation over days as animals become familiar with the test situation. This was clearly not the finding in the present study. The result is not consonant with previous data on inbred strains (Collins, 1966) and may indicate an important difference between inbred and heterogeneous samples.

The utility of multiple crosses as genetic tools

One purpose of producing the eight-way cross was to produce a sample which would resemble more closely the heterogeneity found in wild forms of *Mus musculus*. However, Bruell (1967) has pointed out that the origin of most, if not all, inbred lines commonly used is not the wild *or* commensal form of *Mus musculus*, but more probably domesticated forms of this species. These had probably been in captivity for a long time before being used for inbreeding. Thus the foundation stocks of most inbred strains in existence today may have undergone considerable genetic modification, compared with wild forms, before formal inbreeding was instituted. Thus any attempt to regenerate the nature and content of the original gene pool could only approach that represented by the domesticated mouse. Henderson (1967) has suggested that sampling of hybrids or multiple crosses may be used to extrapolate a 'population standard' relevant to laboratory animals of a given species. This view of the relevance of experimental data based on the behaviour of multiple crosses seems more reasonable in the light of the considerations discussed above. In addition, the homozygosity of the present cross with regard to the albino locus may restrict the relevance of the results considerably, as this locus has been suggested to have an effect on behaviour (Defries *et al.*, 1966; Wilcock, 1969).

Although these points do restrict the generalizability of the data in the ways indicated, certain factors limit Bruell's (1967) point regarding the derivation of inbred lines presently used in behavioural research. For example, there are probably several routes by which genetic changes which produce domesticated forms can be achieved. Thus, genes involved in processes concerned with domestication possibly differ considerably in the various foundation stocks used to derive the various inbred lines. In this light, if a wide enough sampling of inbred lines with diverse ancestry is achieved, it might be feasible to reconstitute the wild genotype with a greater likelihood of success than Bruell (1967) has suggested.

Any future attempt to construct standards of laboratory species would probably be improved by incorporating different coat-colour genes in the mating system. In addition, an estimate of the degree of heterogeneity in wild populations might be used as a basis for selecting the number and relationship of strains used in the cross.

This investigation was supported by NIH Training Grant CA 05013 from the National Cancer Institute, and MH 11327 from the National Institute of Mental Health. I thank Dr J. L. Fuller

and Dr R. L. Collins for useful comments. Mr Frank Clark helped considerably in the execution of the study. Dr E. P. Lee kindly supplied the animals for this investigation and devised the breeding scheme used. He also read the manuscript to my advantage.

REFERENCES

BLIZARD, D. A. (1968). Autonomic and behavioural correlates of emotionality. (Ph.D. thesis, University of Wales.)

BROADHURST, P. L. (1957). Determinants of emotionality in the rat. I. Situational factors. *Br. J. Psychol.* **48**, 1–12.

BROADHURST, P. L. (1960). Experiments in psychogenetics: applications of biometrical genetics to the inheritance of behaviour. In H. J. Eysenck (ed.), *Experiments in Personality*, vol. 1. London: Routledge & Kegan Paul.

BROOKSHIRE, K. H. & RIESER, T. C. (1967). Temporal course of exploratory activity in three inbred strains of mice. *J. comp. physiol. Psychol.* **63**, 549–551.

BRUELL, J. H. (1964). Inheritance of behavioral and physiological characters of mice and the problem of heterosis. *Am. Zool.* **4**, 125–138.

BRUELL, J. H. (1967). Mice and the study of evolution. (Paper presented at the annual meeting of the A.P.A.)

BRUELL, J. H. (1969). Genetics and adaptive significance of emotional defecation in mice. *Ann. N.Y. Acad. Sci.* **159**, 825–830.

CANDLAND, D. K. (1970). The ontogeny of emotional behavior. In H. Moltz (ed.), *The Ontogeny of Vertebrate Behaviour*. New York: Academic Press.

COLLINS, R. L. (1966). What else does the defecation score measure? *Proc. 74th Ann. Conv. A.P.A., 1966*, 147–148.

CROWCROFT, P. (1966). *Mice All Over*. London: Foulis.

DEFRIES, J. C., HEGMANN, J. P. & WEIR, M. W. (1966). Open-field behavior in mice: evidence for a major gene effect mediated by the visual system. *Science, N.Y.* **154**, 1577–1579.

DIXON, L. & DEFRIES, J. C. (1968). Development of open-field behaviour in mice: effects of age and experience. *Dev. Psychobiol.* **1**, 100–107.

FULLER, J. L. & COLLINS, R. L. (1968). Temporal parameters of sensitization for audiogenic seizures in SJL/J mice. *Dev. Psychobiol.* **1**, 185–188.

HALL, C. S. (1934). Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *J. comp. Psychol.* **18**, 385–403.

HALLIDAY, M. S. (1966). Exploration and fear in the rat. *Symp. Zool. Soc. Lond.* no. 18, 45–59.

HENDERSON, N. D. (1967). Prior treatment effects on open field behaviour of mice: a genetic analysis. *Anim. Behav.* **15**, 364–376.

MCCLEARN, G. E. (1960). Strain differences in activity of mice: influence of illumination. *J. comp. physiol. Psychol.* **53**, 142–143.

MCCLEARN, G. E. & MEREDITH, W. (1964). Dimensional analysis of activity and elimination in a genetically heterogeneous group of mice (*Mus musculus*). *Anim. Behav.* **12**, 1–10.

MCREYNOLDS, W. E., WEIR, M. W. & DEFRIES, J. C. (1967). Open-field behavior in mice: effect of test illumination. *Psychon. Sci.* **9**, 277–278.

POILEY, S. A. (1960). A systematic method of breeder rotation for non-inbred laboratory animal colonies. *Proc. Anim. Care Panel* **10**, no. 4, 159.

ROSS, S., NAGY, Z. M., KESSLER, C. & SCOTT, J. P. (1966). Effects of illumination on wall-leaving behavior and activity in three inbred mouse strains. *J. comp. physiol. Psychol.* **62**, 338–340.

WHIMBEY, A. E. & DENENBERG, V. H. (1967). Two independent behavioral dimensions in open-field performance. *J. comp. physiol. Psychol.* **63**, 500–504.

WILCOCK, J. (1969). Gene action and behavior: an evaluation of major gene pleiotropism. *Psychol. Bull.* **72**, 1–29.

WIMER, R. W. & FULLER, J. L. (1966). Patterns of behavior. In E. L. Green (ed.), *Biology of the Laboratory Mouse*, 2nd ed. New York: McGraw-Hill.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

(*Manuscript received* 9 *February* 1970; *revised manuscript received* 18 *May* 1970)

Br. J. Psychol. (1971), 62, 2, pp. 253–260
Printed in Great Britain

THE USE OF QUADRATIC DISCRIMINANT ANALYSIS FOR THE MEASUREMENT OF PROFILE DISTANCE IN SOCIAL PERCEPTION

By A. J. SUTTON

School of Behavioural Sciences, Macquarie University, New South Wales, Australia

The 291 students in the 1965 first-year intake to a variety of degree courses at a London college of technology rated 12 role concepts in each of 35 bipolar, seven-point adjective scales and provided information on their social and cultural background and on certain educational and occupational attitudes.

In many studies similarity between semantic differential profiles in relation to external criteria such as the students' social background has been examined using Osgood's generalized distance index. This index assumes a particular kind of interaction between the profiles and it is the purpose of this paper to demonstrate that several kinds of interaction between profiles can be obtained and that they can effectively be described using quadratic discriminant analysis.

The method of analysis is used to analyse hypotheses arising in the present study. These results are reported and a number of problems concerning interpretation of the results are discussed.

In some recent applications of the semantic differential technique (Osgood *et al.*, 1957) to the measurement of a person's perception of roles which he occupies or with which he comes in contact, a global profile distance index has been used to measure the psychological distance between profiles of role perceptions (Talbot *et al.*, 1961; Hooper & Padden, 1965).

The index was proposed by Osgood *et al.* for the measurement of distance between concepts rated on a set of semantic differential scales.

It is given by

$$D^2 = \sum_{i=1}^{i=n} (x_1 - y_1)^2, \tag{1}$$

where x_1 = rating of first concept of a pair on the ith adjective scale; y_1 = rating of second concept of pair on the ith adjective scale; and n = number of adjective scales.

However, it appears quite inadequate as a measure of the complexity of interaction which could occur between profiles of role perceptions, and its interpretation as a simple measure of psychological distance between a person's perceptions of different roles seems unwarranted. The strongest criticisms of the use of Osgood's global profile distance index have come in a series of papers by Cronbach & Gleser (1953) and Cronbach (1955, 1958). Although they are chiefly concerned with the application of Osgood's D index in the field of interpersonal perception there are no psychometric reasons why their criticisms should not also apply to role perceptions.

In Cronbach's 1958 paper he points out that Osgood's D score describes only one type of interaction between the distributions of ratings of each role concept of the pair. As an example, one hypothesis from the original study (Sutton, 1967) will be

reported in this paper, i.e. that an independent measure of Independence from Family (FI) is related to the interaction between ratings of the concepts Self and Mother for males in a sample of student subjects. If the Osgood D index were to be used in the analysis the hypothesis would have been phrased in 'distance' terms, i.e. that FI would be positively correlated with distance between Self and Mother ratings. In Fig. 1 the distributions of two groups, one of High FI and one of Low FI, are plotted on the one set of axes representing Self and Mother ratings.

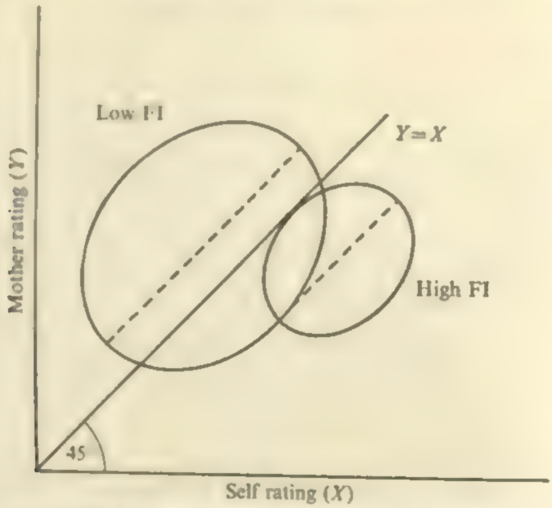


Fig. 1. Distributions of high and low FI groups on a single adjective scale.

The contribution of each point in a distribution to Osgood's D score varies with the length of the perpendicular from the line $Y = X$ to that point. The 'distance' version of the hypothesis stated previously is only supported if the High and Low FI distributions differ in mean distance from the line $Y = X$. Yet there are many ways in which the two distributions might differ without differing in mean distance from $Y = X$. For instance, as Cronbach (1958) points out, they may differ in means, in standard deviations and in slope of regression lines. These parameters would each need to be examined separately in a study of the interaction between Self and Mother ratings in relation to differences in FI.

The univariate procedures required for such an examination are, in effect, carried out by the one process of finding a discriminant boundary which most efficiently separates the two groups, i.e. with the minimum number of misclassifications. In Fig. 1 $Y = X$ approximates such a boundary. However, linear discriminant analysis would fail to discriminate groups which differed only in the variance of the Mother ratings, or in the slope of the regression line. Cronbach (1958) suggests that a complete description of the nature of the different interaction in the criterion groups can be obtained by a quadratic discriminant function, i.e.

$$Z = aX + bY + cX^2 + dY^2 + eXY, \tag{2}$$

where a, b, c, d and e are weights to be determined so that the ratio of between-group variance to within-group variance is a maximum.

Method

The study was concerned with the role perceptions of 291 students of the 1965 intake to the first-year full-time courses in an outer London college of technology. The students were enrolled in pre-degree or degree level courses in arts, sociology, mathematics, business studies and technology. They were asked at the beginning of the academic year to rate 12 role concepts relevant to their position as tertiary students on each of 35 bipolar adjective scales. Data on a number of other sociological and educational variables were obtained at the same time. A portion of this material was collected again from the same students at the end of the 1965–6 academic year. The psychological and educational aims underlying the study are reported by Sutton (1967). Only a portion of the analysis is relevant to this paper. The relevant portion of the original study concerned the relationship between an independently measured psychological variable and the interaction between profiles of ratings of two role concepts, one of which was always Self. The use of quadratic discriminant analysis involves treating separately the ratings from separate scales or factor scores derived from a combination of scales.

Separate scales were used rather than factor scores in the analyses undertaken. This procedure was thought preferable because it was found that the factor structure of the between-scale correlation matrices across all concepts varied somewhat between different subgroups of the total sample. Because the computation of quadratic discriminant analyses for each of the 35 scales for each hypothesis tested would have required more computer time than was available, 14 scales were initially selected to test each hypothesis. They were the two scales which had the highest loadings on each of the seven main factors obtained in an eight-factor maximum-likelihood factor analysis of the between-scale correlation matrix over all concepts and subjects. The scales are shown in Table 1 together with their factor loadings.

Table 1. *Adjective scales initially selected for quadratic discriminant analysis*

Scale	Loading	Factor of highest loading	Percentage variance
1. Pleasant–unpleasant	0·677	1	22·21
2. Friendly–unfriendly	0·692		
3. Obedient–rebellious	0·643	2	9·25
4. Submissive–dominant	0·512		
5. Poor–rich	0·700	3	4·15
6. Low social status–high social status	0·699		
7. Static–dynamic	−0·628	4	2·74
8. Active–passive	0·642		
9. Wise–foolish	0·534	5	1·69
10. Mature–youthful	0·451		
11. Hopeful–disillusioned	−0·462	6	1·30
12. Bored–interested	0·434		
13. Dishonest–honest	0·723	7	1·27
14. Unfair–fair	0·689		

Results

Estimation of the position of boundary in quadratic discriminant analysis

In classification using linear discriminant analysis a discriminant function is obtained and the profile scores of an individual to be classified are substituted in the discriminant function. The individual is classified in the appropriate group according to whether the resulting score is greater or less than the mean of the mean values of each group on the discriminant function. This mean is called the boundary value.

If group sizes are unequal the boundary value is corrected by a quantity equal to the natural logarithm of the ratio of group sizes. These procedures are described in most texts of multivariate analysis, e.g. Kendall & Stuart (1966), Morrison (1967), Anderson (1958). The determination of the boundary which separates groups discriminated on a quadratic discriminant function is not described by these authors. The problem can best be outlined by reference to Fig. 2.

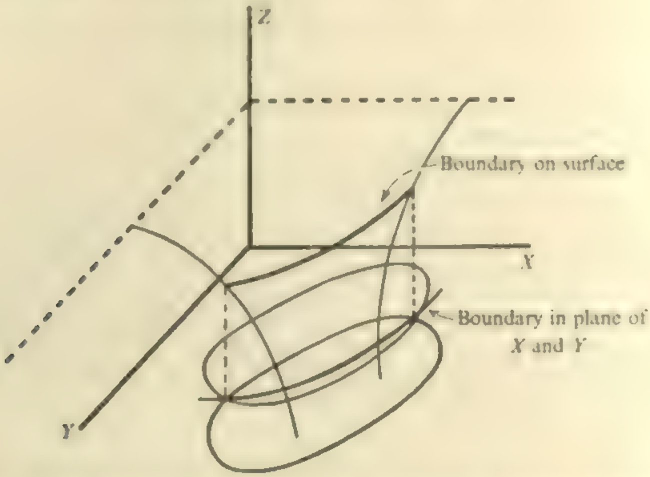


Fig. 2. Three dimensional representation of a quadratic discriminant function.

In three dimensions the quadratic discriminant function represents a surface upon which, if the points of the two groups to be discriminated were projected, they would show maximum separation. The discriminating boundary on the surface passes through the midpoint of the curve joining the mean values on the discriminant function of the two groups. Its projection on the XY plane is a conic section, as shown in Fig. 2.

A precise determination of the position of the quadratic discriminant boundary is not required with the present data. Although the curve shown in Fig. 2 is smooth and continuous, in practice the obtained quadratic discriminant boundary must be placed on a 7×7 matrix representing all possible ratings of Self and Mother concepts. In the present study the matrix of all possible values of each significant discriminant function obtained was calculated. The number of misclassifications was investigated in several cases by imposing the discriminant boundary on the bivariate scattergram of frequencies of each category of response to the pair of concepts. An example is given in Table 2, which shows the bivariate frequency distribution of High and Low FI subjects' ratings on the scale hopeful–disillusioned of the concepts Self and Mother.

It can be seen that the number of misclassifications is high (43 out of 137), despite the high level of significance of the discriminant function. Using Bartlett's test of significance on the first eigenvalue, $\chi^2 = 25{\cdot}29$, d.f. $= 6$, $P < 0{\cdot}001$.

Furthermore, a slight alteration of the boundary does not greatly affect the number of misclassifications. In view of this finding, which was repeated in a number of similar

results, it was considered unnecessary to proceed with the precise placing of the quadratic discriminant boundary. Instead, the mean of the group mean discriminant function values, corrected for unequal group sizes, was taken as a satisfactory approximation of the boundary value.

Table 2 *Bivariate frequency distribution of ratings of Mother and Self concepts on scale hopeful–disillusioned*

High group frequencies are in the upper left corner of each cell, low group frequencies in the lower right corner. The quadratic discriminant boundary is shown as a heavy line: the region on the left of the heavy line (quadratic discriminant boundary) represents those categories of response indicating most probable Low group membership and the region on the right most probable High group membership.

Mother \ Self	1	2	3	4	5	6	7
7		2					1
6	1	2 / 1	2		2	1 (low)	
5	1 / 1	3 / 1	3	1		1	
4	1 / 2	5 / 2		1	1 (low)		1 (low)
3	2 / 1	4 / 2	6 / 2	1 / 1	1		
2	4 / 7	11 / 13	1 / 2	3 / 3	1 / 1	1	2
1	6 / 16	3 (low)	3 (low)		1		3

Interpretation of results

To fully investigate the hypothesis that subjects' scores on an independent measure of Independence from Family (FI) are related to the interaction between Self and Mother ratings, it was necessary to calculate quadratic discriminant functions for each of the 14 scales listed in Table 1. The groups to be discriminated were designated High, Medium and Low FI and for 10 of the scales a quadratic function of the concept ratings discriminated between two or more of the three groups. The discriminant boundaries were each sketched on the 7×7 matrix of possible ratings.

It proved a complex task to summarize the results obtained (there were 11 other hypotheses similar to the one illustrated here). A summary scheme was adopted in which the boundaries were interpreted according to their similarity to a number of

types of linear boundary (i.e. according to their approximate slope). These are shown in Fig. 3.

In both Type 1*a* and Type 1*b* the discrimination is due to variation in only one of the concepts, Self in Type 1*a* and Mother in Type 1*b*. Interaction between the concepts of a 'distance' kind, as discussed previously, is shown by Type 2 and interaction which discriminates the groups according to the relative size of the ratings of both Self and Mother together is shown by Type 3.

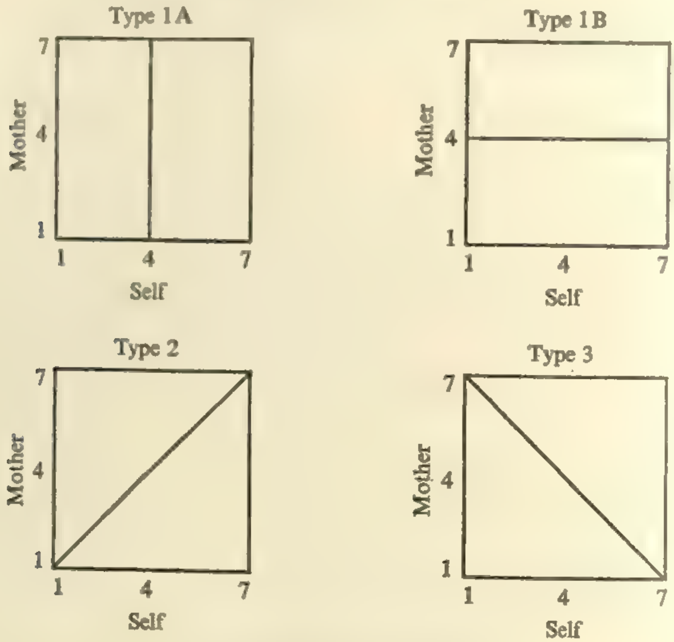


Fig. 3. Types of interaction.

Seven out of the ten significant discriminations could be classified in Type 3. A typical example is that of the scale hopeful–disillusioned which is illustrated in Table 2. The scales on which the groups were discriminated in this way were:

unfair–fair	mature–youthful
friendly–unfriendly	hopeful–disillusioned
wise–foolish	pleasant–unpleasant

low social status–high social status

For ratings on these scales the High FI group is distinguished from the Low FI group and sometimes the Medium FI group by a less favourable rating of both Self and Mother. This type of interaction is not of the 'distance' kind and these results show that the use of a 'distance' index would have failed to reveal the major kind of interaction between Self and Mother ratings when discriminating FI groups.

Two of the significant discriminations were of Type 1*a*, where the discrimination is chiefly due to variation of ratings of Self. These two scales were obedient–rebellious and poor–rich, for each of which the discriminant function was a vertical straight line.

In both cases High FI students tend to have a less favourable opinion of Self than Low FI students.

The remaining significant discrimination was for the scale active–passive. The discriminant boundary is shown in the following figure:

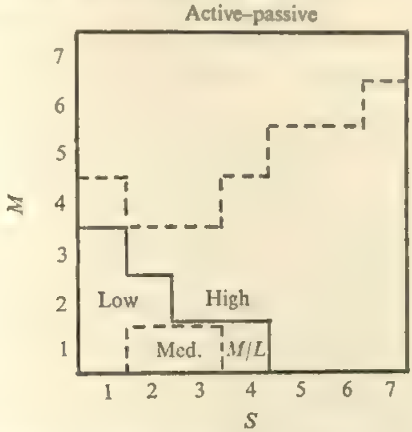


Fig. 4. Discriminant boundaries of scale active–passive. Two discriminant functions were required. High FI was discriminated from Medium FI on function I (——) and Medium FI was discriminated from Low FI on the orthogonal function II (– – –).

The diagram in Fig. 4 is typical of a number of results which could not be interpreted. Interpretation is particularly difficult when two discriminant functions are required to discriminate the groups, or when the one discriminant function has different shapes for each boundary value.

Discussion

The use of quadratic discriminant analysis to investigate interaction between ratings of role concepts has shown that types of interaction that are not measured by the 'distance' index can be significantly related to a criterion variable.

In the other hypotheses investigated in this study similar results were obtained to the illustrative results described above. In this sense, the suggestion by Cronbach (1958) for the use of quadratic discriminant analysis in such cases has proved useful. However, as a practical method of analysis it has disadvantages.

When investigation of the interaction of a pair of concepts in relation to a criterion variable involves discriminations for each of several dimensions, the method is very time-consuming. Rather than obtain discriminant functions for each of a large number of separate scales, it would be better to use dimensions that are the result of considerable prior investigation – for instance, Evaluation, Activity and Potency as in Osgood's original work. If, as in the present case, such dimensions are less stable with social concepts, some attempt should be made to establish stable and meaningful dimensions for the social situation under investigation.

As shown in Fig. 4, the use of three groups may result in two discriminant functions or in different shapes of the one function for two boundaries. The interaction is then difficult to interpret and the most obvious solution is to use only two groups in each discriminant analysis.

Quadratic discriminant boundaries were placed in the diagrams illustrated in the

above figures by an appropriate method. As pointed out previously, a complete solution did not seem worth while in view of the short range of the rating scales and consequent high number of misclassifications in the discriminations. Although Osgood et al. (1957) have argued that seven-point scales are of optimum length with the semantic differential technique, it appears that a thorough use of quadratic discriminant analysis in the present problem would require longer scales, say of 11 divisions. Gulliksen (1958) has also proposed lengthening the semantic differential scales.

It may appear from the above comments and from the necessity to interpret the data by its approximation to certain linear boundaries that linear discrimination would have been adequate in the analysis of results in the present study. The finding that most of the significant discriminations were of Type 3 would have been obtained using linear discrimination but this could not have been known beforehand. In another set of data the discriminant boundaries might, for example, have been consistently hyperbolic and orientated in a particular way. A different classification of boundaries for interpretation would then have been required.

As Cronbach (1958) suggested, quadratic discriminant analysis has been useful in investigating the interaction of ratings of pairs of concepts in relation to a criterion variate. As with interpersonal perception, the use of oversimple profile distance indices in social perception is often a hindrance to adequate interpretation of results, not a help. The complexities of social interaction require procedures of generality, and quadratic discriminant analysis seems a useful technique in this area.

The research reported in this paper was conducted while the author was Research Fellow at the Enfield College of Technology, Enfield, Middlesex, and support for the project by the London Borough of Enfield and the College authorities is gratefully acknowledged. It forms part of a thesis accepted for the degree of Ph.D. by the University of London and the author is grateful for the encouragement and advice of his supervisor, Professor P. E. Vernon.

REFERENCES

ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. New York: Wiley.

CRONBACH, L. J. (1955). Processes affecting scores on 'understanding of others' and 'assumed similarity'. *Psychol. Bull.* **52**, 177–193.

CRONBACH, L. J. (1958). Proposals leading to analytic treatment of social perception scores. In R. TAGIURI & L. PETRULLO (eds.), *Person Perception and Interpersonal Behaviour*. Stanford, Calif.: Stanford University Press.

CRONBACH, L. J. & GLESER, G. C. (1953). Assessing similarity between profiles. *Psychol. Bull.* **50**, 456–473.

GULLIKSEN, H. (1958). How to make meaning more meaningful. *Contemp. Psychol.* **3**, 115–119.

HOOPER, D. & [illegible], D. (1965). Psychiatric roles and their meaning. *Br. J. soc. clin. Psychol.* **4**, [illegible]–38.

KENDALL, M. G. & STUART, A. (1966). *The Advanced Theory of Statistics*, vol. 3. London: Griffin.

MORRISON, D. G. (1967). *Multivariate Statistical Methods*. New York: McGraw-Hill.

OSGOOD, C. E., SUCI, G. J. & TANNENBAUM, P. H. (1957). *The Measurement of Meaning*. Urbana, Ill.: University of Illinois Press.

SUTTON, A. J. (1967). The measurement of students' roles using the semantic differential technique, and their relation to social background and satisfaction with studies. (Unpublished Ph.D. thesis, University of London.)

TALBOT, E., MILLER, S. C. & WHITE, R. B. (1961). Some aspects of self conceptions and role demands in a therapeutic community. *J. abnorm. soc. Psychol.* **63**, 338–345.

(*Manuscript received* 29 *November* 1968; *revised manuscript received* 20 *March* 1970)

Printed in Great Britain

PAPERS TO APPEAR IN FORTHCOMING ISSUES

(Not previously listed)

[illegible], W. G. & C[illegible], H. H., Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pa., U.S.A. [illegible] in the perception of [illegible].

[illegible], P. R., Department of Psychology, West Ham College of Technology, Stratford, London. [illegible] aspects of negation in sentence evaluation and comparison tasks.

D[illegible] & A[illegible], E. C., Department of Psychology, American University of Beirut, Lebanon. Some further observations on the measurement of clustering in free recall.

Davies, J. B., Department of Psychology, University of Strathclyde, Glasgow. New tests of musical aptitude.

[illegible], K. D., Department of Psychology, University of Hull. Long-term retention and transfer of an industrial search skill.

Frith, U., M.R.C. Developmental Psychology Unit, Gordon St, London W.C.1. Why do children reverse letters?

Gale, M. A., Coles, M. G. H., Kline, P. & Penfold, V. L., Department of Psychology, University of Exeter. Extraversion–introversion, neuroticism and the EEG: basal and response measures during habituation of the orienting response.

Harris, P. L., Institute of Experimental Psychology, University of Oxford. Examination and search in infants.

Innes, J. M., Department of Psychology, University of Birmingham. Word association, associative structure and manifest anxiety.

[illegible], W. J., Department of Philosophy, University of Leeds. On the terms 'subliminal perception' and 'subception'.

Br. J. Psychol. (1971), 62, 2, pp. 263–288
Printed in Great Britain

BOOK REVIEWS

Conditioned Reinforcement. Edited by D. P. HENDRY. Homewood, Ill.: Dorsey Press. 1969. Pp. xxii+454. £4.40.

This book is the result of a symposium on conditioned reinforcement which was held in 1967. Conditioned (or secondary) reinforcement is a concept which is mentioned in passing in most textbooks of psychology, though rarely with much conviction. It is clear that such a concept is necessary to explain behaviour within reinforcement theory, for animals and humans cannot be said to gear their behaviour only to the traditional primary reinforcers of food, water and sex. So has arisen the idea that a stimulus may become an acquired reinforcer by reason of an animal's past history. On the other hand, experimental investigations of such conditioned reinforcement have not been particularly impressive – mainly because the situations traditionally used have guaranteed transitory demonstrations of the phenomenon (for example, findings that extinction of a pattern of behaviour can be retarded if conditioned reinforcers are still delivered when primary reinforcers are withheld).

Two hypotheses of the nature of conditioned reinforcement have been influential. The first, attributable to Hull, is that a stimulus acquires the characteristics of a secondary reinforcer by occurring in close temporal contiguity with a primary reinforcer. The second, developed from Skinner, is that a stimulus becomes a conditioned reinforcer if it sets the occasion for a pattern of behaviour which obtains primary reinforcement.

Hendry's book presents a great deal of data, which result in the main from experiments which would be described as 'operant conditioning'. In the hope, presumably, of making these data more intelligible to the non-specialist, Hendry opens by providing a simple discussion of the characteristics of such experiments, under such headings as 'What is an operant?' and 'Study of individual organisms'. After these preliminaries, there follow an impressive number of reports demonstrating the control of behaviour by events which have not traditionally been regarded as reinforcing in their own right. Of special note is the inclusion of data which have been extremely influential in this field although not formally published hitherto (Autor, ch. 6; Wyckoff, ch. 9). The first few reports explicitly demonstrate the prolonged maintenance of patterns of behaviour by conditioned reinforcers. This is achieved by making them contingent upon behaviour according to some schedule, while primary reinforcers are programmed by a separate schedule. Such prolonged maintenance of behavioural control may have more apparent relevance to the assumed effects of conditioned reinforcement outside the conditioning laboratory than the earlier extinction tests. The next section contains interesting reports of animals' preferences for different combinations of schedules and stimuli as these are affected by conditioned reinforcement. Animals are allowed to choose between such combinations by means of concurrent chain procedures, a refinement in experimental analysis which may offer considerable advantages in this area of research. The remaining chapters are marked by a linking theme in which the concept of information plays a dominant role. This is an extension of an important contribution to the theoretical analysis of conditioned reinforcement made by Egger and N. Miller. They proposed that a stimulus becomes a conditioned reinforcer to the extent that it provides information about the occurrence of primary reinforcement. This idea receives much attention in the later chapters of Hendry's book, and seems to be generally favoured on the basis of experimental manipulations which are considerably more sophisticated than the extinction tests employed by Egger and Miller.

The data presented in this book establish a number of findings which are important from a theoretical point of view. In particular, it is clear that direct temporal pairing with a primary reinforcer is not necessary for a previously neutral stimulus to acquire conditioned reinforcing properties (e.g. Kendall's ch. 10 on the reinforcing properties of trace stimuli). Indeed, stimuli which can *never* be associated with food may acquire conditioned reinforcing properties in that they will control patterns of behaviour appropriate to the ways in which they are scheduled (Marr, ch. 2; De Lorge, ch. 3). One implication of these findings may be that conditioned reinforcers cannot be regarded as effective solely because they are paired with need reduction,

for example. This, of course, detracts from the suggestion that conditioned reinforcers acquire some of the motivational properties of a primary reinforcer.

It is clear from the book that this is at present an extremely active area of research, and although many questions are posed by these data, not all the answers are currently available. Hendry's conclusion is that the effectiveness of a stimulus as a conditioned reinforcer 'may be determined partly by its association and partly by the information it conveys' (p. 400). This introduction of the concept of information to operant conditioning is greatly to be welcomed. The quantification of the environment offered by information theory, for example, should increase the effectiveness of theories which confine themselves to examining the interactions between behaviour and the environment. One interesting feature, however, is that the concept of information may easily lead to the concept of active information-processing by animals. How far this can be avoided by Skinnerian analyses of behaviour is an intriguing question. DEREK BLACKMAN

Body Experience in Fantasy and Behavior. By SEYMOUR FISHER. New York: Appleton-Century-Crofts. 1970. Pp. xii + 690. $13.75.

'The experienced body is a world within a world. It is a complexly shaped yet bilaterally symmetrically simplified object which has spatial and geographical features second to none in their diversity.'

Seymour Fisher then sees the body world as parallel to the environmental one, and the study of body experience as a partner to it in personality study. Although body experience has not perhaps been exactly a sleeping partner, it has until recently been relatively quiescent. In this volume Fisher documents the recent researches of himself and, in the main, his students. It is a continuation of the ideas and studies reported with Cleveland in *Body Image and Personality*. The related work of Witkin, Secord, Jourard, Werner and Wapner is considered essentially only in so far as it bears on the Fisher findings.

In body study, as in other personality research, workers differ in the dimensions of variation or factors that they find important and predictive, and they differ in the techniques of measurement. The particular Fisherian variables are Body Prominence, Body Boundary, and Body Landmarks, which are the parts that we particularly attend to. Theoretically he sees the body and its parts as a system for guiding behaviour and evaluating alternatives. The background for this view is that all of a child's early judgements about the environment are based on body states – the body tells what is 'good' and what is 'bad'. This view is, of course, acknowledged to be dissonant with current 'central' theories of cognition and emotion. However, it is consonant with basic psychoanalytic notions and with certain fundamentals of Piaget. At least it seems reasonable to agree with Fisher's proposition that body feelings provide one way of evaluating one's experience, and that habitual modes may be found.

The three basic measures used are Holtzman-type Rorschach testing for barrier scores, the Body Focus Questionnaire (in which the subject checks one of a series of two alternative parts that he is more aware of), and a spontaneous listing of 'The twenty things you are aware of at this moment' which is scored for Body Prominence.

The validity of these indexes of body experience is then tested in terms of their predictive power with respect to other measures of body experience, personality traits, social background variables, and perceptual and learning functions. These cover a catholic range: content of first memories, tachistoscopic perception, word association, performance on the Thereness-Thatness apparatus, appreciation of jokes, among them. Predictions are made often on the basis of psychodynamic interpretations, and result sometimes in rather startling combinations – 'mouth and stomach awareness in women with power orientation'; 'heart awareness in men with religiosity and guilt'.

There is a plethora of data, some of them puzzling even to Fisher in their meaning. However, there does appear to be a core of stable and replicated findings that are of general interest. Women have higher barrier and lower penetration scores than men; they evidence a more constructive flexibility in their coping with changed body stimuli. Men show particular correlations between awareness of the right side of the body and heterosexual difficulties. Perhaps the most impressive of the empirical relationships documented is that between highly developed boundary scores and good responses to stress. The reviewer would be less sure than Fisher that boundary is the causal variable, but it does seem certain now that there is a connexion between

high barrier scores and constructive adjustment to bodily disablement, due to amputation, poliomyelitis or paraplegia. Testing the limits of the clinical finding, implications for other stress situations are explored, and evidence suggests that manifest anxiety in women undergoing gynaecological examination, low pain tolerance in a psychological experiment, and low efficiency in dealing with conflict stress in a cognitive task, are all related to low barrier scores.

One point in Fisher's series of studies deserves special mention as a promising innovation in this genre. This is the introduction of the experimental method. If natural individual differences have certain personality correlates, will the production of heightened differences between an experimental and control group confirm the correlation? The strategy used is to focus special attention on particular regions (by physical stimulation or by asking the subject to monitor his own heart rate, for example) and then to test for heightened scores in certain normally correlated functions. Although the results are not uniformly supportive, this is a healthy direction of interest.

The exploration of body-image geography and ecology is not yet systematized, but it has an intrepid explorer here, and the book is an intriguing reference for anyone interested in the whole man.

HALLA BELOFF

Information and Choice. By JOHN COHEN and IAN CHRISTENSEN. Edinburgh: Oliver & Boyd. 1970. Pp. vii+148. 38p.

At a time when synthesis rather than analysis in psychology is most welcome, the authors of *Information and Choice* set themselves a formidable task. They attempt to indicate some of the relationships between subjective aspects of risk and decision making, strategies in search behaviour, and interpretations of probability and information.

In their efforts to develop the synthesis, they focus on definitional problems. Cohen and Christensen acknowledge that they 'do not pretend to have achieved anything like the necessary logical coherence this family of concepts deserves' (p. 142). However, to suggest that there may be relationships among their chosen set of concepts may be all that can be expected in 150 pages.

The organization of the book is a description of each concept in turn. This is perhaps an unfortunate format for a book emphasizing dependence among concepts. Each chapter, with a few exceptions, focuses on the definitional and logical problems surrounding its theme. The first, on information theory, presents some background and an emphasis on subjective as opposed to objective information. This is followed by a chapter on probability, similarly highlighting subjective probability. Chapters on deciding and choosing, risk-taking, and search behaviour ensue, with some useful criticism of variables often defined differently by experimenters working on the 'same' problem.

Exceptions to this general emphasis on definition are to be found in the chapter on search behaviour, and in the appendices, e.g. probability and the law, a discussion of jury members' feelings of doubt. Experiments performed by Cohen, in collaboration with Meudell, are reported in the search behaviour chapter. For experimental psychologists, especially those unfamiliar with the subjective probability literature, this would seem to be the most convincing chapter.

The presence of much vital information in the book in footnotes or appendices, and irrelevant information in the body of the writing, e.g. translations of words or phrases into continental tongues, is often irritating to the reader. Perhaps these aspects of the book are the result of the authors' attempt to write comprehensively as well as originally for 'scientists or housewives, politicians or painters, soldiers or theologians' (back-cover). On the other hand, the addition of occasional examples from the theatre and current affairs as well as from experimental work is often refreshing.

The book seems uneven in emphasis and direction. The experimentalist would probably prefer more data, the philosopher more rigour, and the housewife more examples. As mentioned before, this writer feels this variability may be due to the difficult objective Cohen and Christensen chose for themselves: writing with wide scope and purpose for a diverse readership. However, these aspects of the book do not detract from the psychologically interesting task of specifically relating the concepts of probability, risk, decision-making, and search, across areas of inquiry, and indeed within areas.

C. WEIR

Attention and Performance. III. Proceedings of a Symposium on Attention and Performance. Edited by A. F. SANDERS. Amsterdam: North-Holland. 1970. Pp. vi + 442. £8.20.

This book is the same as volume 33 of *Acta Psychologica*. It is a sequel to 'Attention and Performance I', volume 27 of *Acta Psychologica*. ('Attention and Performance II', volume 30 of *Acta Psychologica*, is devoted to reaction times.) The secret of Andries Sanders's success is that he invites to his symposia the world authorities on attention and performance. The result is a series of books which are a must for every department of psychology which claims to be concerned with cognition. In this volume III about two-thirds of the papers describe original work which has not been published before, although few of the papers bring out anything which is startlingly new. The remaining papers present models, general discussions, and summaries of work published elsewhere. As in the previous volumes, the papers are divided into a number of topics. 'Selective attention' and 'Subjective probability' appear for the first time in the series as titles for groups of papers. The editor supplies a brief introduction which is printed before each group of papers.

In a collection of papers, those selected for special mention represent as much the interests of the reviewer as the quality of the particular papers. There was a novel condition in one of Anne Treisman's experiments on dichotic listening. She presented nonsense syllables like TAV and SEM simultaneously, one to each ear. When the listener is asked to report either one of them, he may report TEM or TEV. The former involves one switch between the two ears, the latter involves two switches. Anne Treisman found an average of 30 per cent of switches between the pairs of nonsense syllables. A similar percentage was found when the pairs of nonsense syllables were presented binaurally to both ears. Thus in dichotic listening the inputs to the two ears are not perfectly tagged in the acoustic store. The separation is better if a digit is presented with the nonsense syllable. Presumably the listener recognizes the digit, and then reports the remainder of what he has heard as the nonsense syllable. This is not so easy to do when the digit and nonsense syllable are presented binaurally, owing to the masking of each by the other. Thus the major advantage of dichotic presentation is in better reception, or reduced masking. Dichotic presentation does not always put the inputs to the two ears in recognizably different locations in the acoustic store.

John von Wright has been experimenting on the short-term visual storage of coloured letters. He presented by tachistoscope eight X's and eight O's, or eight E's and eight F's, arranged randomly in a 4 × 4 matrix. Four of each kind of letter were red, the other four blue. In some trials the man was asked directly after presentation to mark the positions and colours of a particular letter. In other trials he had to mark the positions and letters of a particular colour. With the X's and O's, which are easy to discriminate, reports by letter and by colour were equally accurate. But with the E's and F's, which are difficult to discriminate, reports by letter were less accurate than reports by colour. During the time which it took to pick out the E's from the F's, the colours of some of the E's, or perhaps their positions, had faded beyond recall. In this kind of experiment, too many items are presented. To recall as many items as possible correctly, the man has to be able to discard the unwanted items as quickly as possible. Increasing the difficulty of the discrimination which has to be made before items can be discarded has a similar effect to delaying the instruction on what can be discarded.

New experiments are also reported on the physiological correlates in the brain of expectancy or vigilance. When two signals follow each other at a constant interval, the person soon expects the second signal. The cortex develops an expectancy wave ('contingent negative variation') just before the second signal evokes its response. In an auditory vigilance task, Bob Wilkinson found that people miss fewer signals when they show a large expectancy wave at the vertex.

Risto Näätänen recorded from the vertex of the cortex, from the occipital region, and from the temporal region. He presented alternately loud and soft clicks, one click per second. The observer had to listen either to the loud clicks or to the soft clicks, to detect an occasional small change in intensity. Two out of the five observers showed a good deal of enhancement of the evoked responses to the clicks which they attended to, as compared to the clicks which they did not attend to. But they did so both in the vertex and in the temporal region. Also the enhancement was accompanied by a marked expectancy wave at the vertex, and by decreased

amplitude of the EEG in all three regions of the cortex. Risto Naatanen argues that this indicates a non-specific arousal of the whole of the cortex. It does not indicate selective attention to the auditory clicks which had to be attended to. These illustrations represent some of the more novel findings reported in this excellent volume.

E. C. POULTON

Visual Thinking. By RUDOLPH ARNHEIM. London: Faber & Faber. 1970. Pp. xi + 345. £5.50.

Anyone knows what it is to think visually if he paints pictures, produces theatre, does jigsaw puzzles or birdwatches. Teachers use visual patterns to help pupils understand a new complex of relationships. Psychometrics reveal correlates of ability to deal with visual materials. Some people are even convinced that their most ambitious thinking is more visual than verbal and that, as Ribot said, the logic of images is the prime mover of constructive imagination. So visual thinking is important. The notorious difficulty is that it resists verbal dissection. Because of this and because verbal language has had such heavy innings in recent years, we can turn hopefully to this book by a respected writer on visual arts who is also maturely informed about cognitive processes and is now Professor of Psychology of Art at Harvard. He writes quote-worthy prose which sparkles with psychological *aperçu* and he touches securely on an impressive variety of issues involving three great problems, namely, the organization of visual perception, the nature of abstraction, and the relations between verbal and pictorial thinking. We are conducted on a lively tour of visual perception and thinking. We are left in no doubt that visual thinking merits further investigation. We are given a helpful bibliography. And for all this we can properly be grateful. But regarding the ways in which visual thinking actually works, we are told little that we have not already been told by the original gestalt triumvirate—for whom Arnheim has strong sympathy. It would be fair to say that the book excitingly reconnoitres a thicket of vital psychological problems without telling us precisely how these problems are to be solved.

Why this disappointing outcome from a work so rich, comprehensive and wise? There are three possible reasons. Firstly, Arnheim addresses the cultivated reader who may be uninformed about many of the issues and who is presumed to prefer the illuminating commentator to the hard systematizer and experimentalist. Secondly, there are inherent difficulties in translating visual processes into verbal ones and *vice versa*: a film about visual thinking might be more successful than any book could hope to be. Thirdly, the book's theme is wider than visual thinking and is allowed to distract. This theme is the unity of perception and thought. A dichotomy between these two is held to be misleading because all perceiving is itself intelligently concerned to grasp generalities while 'all thinking (not just thinking about art) is basically perceptual in nature'. This theme is scarcely novel. But it is worth stating because the genuine similarities between the activities we call perceiving and thinking are perhaps less often stressed than the equally genuine differences. The trouble is that Arnheim is tempted by the grandeur of this theme into scholarly byways which, though brilliant in themselves, blur rather than clarify substantive issues. He could certainly have given us more if he had aimed to give less.

Among readers of this *Journal* the book, by virtue of style and price, will appeal to only two small groups. One group will enjoy having a perceptive horizon-widening volume for, as it were, the psychological coffee table. The other will have read Eleanor Gibson's *Principles of Perceptual Learning and Development* and Richard Gregory's *The Intelligent Eye* and will be searching for clues on how studies of visual perception might be extended into the realms of thinking, education, aesthetic appreciation and everyday communication. For this group of readers the book provides clues in plenty.

IAN M. L. HUNTER

Approaches to Thought. Edited by JAMES F. VOSS. Columbus, Ohio: Merrill. 1969. Pp. xii + 337. $9.50.

Approaches to Thought is a collection of papers read at a symposium held at the University of Pittsburgh in October 1966. The editor, J. F. Voss, states that the aim of the conference was 'to provide an opportunity for individuals to discuss or speculate on the issue of how their general research interests could be relevant to the study of thinking', the need for such a

discussion being given by a shift in emphasis 'from the externally presented stimulus as the major cause of behaviour to the role of factors within the organism as a primary cause of behaviour'. The principal interest in the collection, therefore, is that it presents, in a single volume, a picture of a group of American psychologists of various persuasions grappling with the relationship between their own theoretical outlooks and an elusive and largely undefined topic, thought. A book of this kind clearly runs a danger of being seriously out of date by the time it is published, unless it can convey the flavour of a genuine debate and includes some programmatic definitions which remain, in Voss's words, 'thought-provoking'.

The debate in Pittsburgh in 1966 seems to have been between a rather defensive neo-behaviourism, represented by papers on verbal learning by Voss and Battig and a postscript to the conference by Kimble, and 'information processing' approaches enthusiastically described by Uhr, Reitman, Simon and others. As such, the book might well have some value as an introduction to contemporary study of cognition, and as a source of points for discussion.

The collection includes a reasonable sample of viewpoints, although there is a conspicuous absence of contributions from differential psychology or developmental psychology. Haber and Posner give summaries of work in the area of visual cognition which depends on Garner, Hake and Eriksen's notion of 'converging operations' as an approach to the experimental definition of non-observable 'processes'. This position is also represented by Newell, who cites the work of Sternberg, and by discussions of mathematical models of search behaviour by Restle and Anderson. Newell's comments are made in reply to a long paper by Bourne, in which Bourne argues against talk of mental processes but in favour of descriptions of behaviour as a rule-governed enterprise. Bourne's position is that 'any particular behaviour (response or response sequence) committed by the organism is recognizably consistent with and instantiates a rule', but that this does not imply 'the existence of underlying and/or antedating psychological processes which are said to control, govern, or regulate responses'. This argument, which clearly owes much to Ryle's *The Concept of Mind*, is seen by Newell as 'almost a failure of nerve' since it outlaws process theories which may generate useful predictions. In fact, the debate here is possibly an empty one, since the term 'process' is used in different ways by Bourne and the 'information processing' advocates. Reitman gives a fairly detailed review of studies using computer programs as representations of models of concept attainment, pattern recognition and grammatical competence. Reitman argues that the use of a computer permits 'precise statements directly in terms of complex structure and process systems, and allows strict tests on the sufficiency and implications of our models'. The same point is stated more emphatically by Uhr ('The computer gives every one of us the power of a million, or a billion, graduate research assistants'), who sees the simulation movement as forcing the precise specification of terms and making possible the development of a genuinely hypothetico-deductive method in psychology.

In summary, Pittsburgh 1966 seems to have been something of a field-day for an approach to the study of thought which accepts the reality of mental processes and seeks to represent the logical structure of such processes, perhaps using a computer to test out the full implications or inadequacies of any particular model. By comparison, the more traditional S-R contributors appear disappointed in the progress they report and doubtful of the relevance of their methodologies to the study of thought.

P. H. K. SEYMOUR

Experimental Psychology: Its Scope and Method. Volume 7. *Intelligence.* Edited by PAUL FRAISSE and JEAN PIAGET. London: Routledge & Kegan Paul. 1969. Pp. ix+283. £1.75.

The seventh in this extensive series of textbooks was first published in French in 1963 and consists of four long chapters by Oléron, by Piaget and Inhelder (2) and by Gréco, entitled respectively 'Intellectual Activities', 'Mental Images', 'Intellectual Operations and their Developments', 'Learning and Intellectual Structures'.

It must be said at once that, although an extremely competently written collection full of interesting insights and comments, this book is hardly suitable as a general textbook on intelligence and thinking for the British or American undergraduate student. Although a wide range of literature, including much American and British research, is reviewed, the treatment remains very firmly within the Piagetian framework. The first, third and fourth chapters (chapters 22,

24 and 25 in the total work) may be described as overlapping variations on several basic themes. These include (*a*) the classification of types of reasoning and of the experimental procedures used to study them, (*b*) the Piagetian analogy between 'operations' in concept formation and 'operations' in logic and mathematics, (*c*) an attempt to integrate conceptually the two kinds of activity (or the two kinds of description common in psychological textbooks) typically grouped under learning and thinking.

The reader will look in vain for any reference to the genetics of intelligence, to environmental or sociological studies of factors affecting the development of intellectual ability, or indeed to the existence of individual differences in human ability. Surprisingly, the only reference to individual differences of this kind appears to be to differences in the giftedness of guinea-pigs reported by Bartlett in 1932.

The vocabulary is often esoteric, though this may be partly the result of translation; the student will surely have trouble with some of the following, since they are all introduced without direct explanation – criteriology, syncretism, hyperordinal anticipation, stereognosis, logicism, parasitic elements and isotropic premises.

To consider the four chapters briefly in turn, Oléron on 'intellectual activities' comes nearest (for good or ill) to the accepted style of a textbook, making stern efforts to classify the nearly unclassifiable data in the realm of thinking and problem-solving. Piaget and Inhelder offer a very interesting addition to the scattered and until recently beyond-the-pale literature on mental images. They also, in their other chapter, provide one of the best summaries, straight from the horse's mouth, of 'intellectual operations and their development'. Piaget's views on such difficult topics as 'equilibration' and 'the relation of language to thought' are clearly and briefly summarized. Finally, Gréco ably traces similarities and differences between the 'learning theory' and the Piagetian approaches to human learning.

The book is well produced, reasonably priced, and should be recommended reading for the advanced student in this field.

H. J. BUTCHER

On the Biology of Learning. Edited by KARL H. PRIBRAM. New York: Harcourt, Brace & World. 1969. Pp. vi + 225. £2.15.

The five essays printed in this book were first given as a series of public lectures sponsored by the Graduate School of Education of Harvard University during 1966–7. The brief given to each speaker was to anticipate and expound issues in the fields of the neurophysiology and biochemistry of learning and to relate them to the wider aspects of education.

Whilst reading this book the immediate impression is of discontinuity, as the essays are not complementary; however, the reviewer's final impression was that the essays did serve to make a rounded review. Certainly they make interesting reading for anyone teaching psychology from the aspect of biological antecedents. And for any educationist who wished to understand the biological approach to the study of behaviour they will be of immense value.

Lorenz's essay takes up nearly one third of the total length of the book. In any short review it is always difficult to present a fair and balanced view of Lorenz's writing. His range is huge, his points and theories very palatable, and it is only in retrospect that one begins to doubt. The flavour of this essay is best described as Lorenz *à la* Koestler; for bisociation, now read fulguration. From my profuse notes I was hard put to improve upon Karl Pribram's pithy editorial remarks. Commenting upon Lorenz's contribution he writes, 'examples are perhaps coloured, evidence too often preliminary and nonquantitive to give assurance...However, as with any vision expressed, a certain one-sidedness is unavoidable.' I feel that Lorenz fails to achieve the synthesis between explanations of biological and social evolution. But his attempt stands as a salutary answer to those who deny that academics ever attempt such an undertaking.

Hyden's essay, which deals with macromolecules and learning, after the most cursory background introduction plunges into details of his beautiful and skilful methods. To those for whom phase-contrast microscopy, X-ray microspectrography and electrophoresis are mere biological incantations, little will be gained. However, this chapter is a rich source of information for those knowledgeable about this area. One interesting point which Hyden raises relates to the experiments purporting to show inter-animal transfer of learning via extracts of macromolecules. Hyden claims that the transfer of behaviour via brain extracts is a phenomenon that exists, but that we still lack understanding of the actual mechanism.

Magoun, also, deals in part with macromolecules and learning, and he presents some of the theories for and against the macromolecular approach. He warns, as does Pribram, against a too naive psychobiological approach and also of making simplistic biological analogues for educational programmes. A good example is the American Headstart educational programme. This approach appears to have assumed that, like a vitamin deficiency, all that was required for backward children was an educationally enriched environment. The problem, however, has proved to be a great deal more complex and intractable. Starting from John Locke's essay *Of Ideas*, Magoun considers some of the attempts to provide psychobiological explanations of learning at various physiological levels. Pointing to the discrepancy between Sokolov's neuronal model, which stresses orientation to novelty, and repetition, the so-called primary law of learning, Magoun adds an interesting point to the current questioning of the concepts relating to reinforcement by stating that the value of reinforcement is that it prevents habituation.

Penfield's essay is an excellent summary of his work. His section on second-language learning, subtitled 'an aside to educators', will probably go unnoticed by the majority of them. This will be a great pity as his childhood time-table for language learning and his statements about the bilingual child and adult, dealt with in a most illuminating way, would alone be worth the price of the book.

Pribram, apart from his excellent introduction to the book, deals in the first part of his essay with the technique and principles of holography and explains how they can be used in a theory of neural holography (holography, or an alternative explanation of cross-correlograms, has recently been used to explain how coded information is stored in the brain, taking into account the survival of such information following cerebral damage). He emphasizes the bio-behavioural role of educators and asks them to remember that they are educating the brain. This point might also be extended to cover teachers of child psychology courses; all too often the children they talk about are apparently without brains and have only rudimentary, if any, sensorimotor development. Using the holographic analogy Pribram deals with the four R's of remembering, which are stated as: representation, registration, reconstruction and rearrangement. He concludes that consideration of such a theory indicates that educators have neglected the total image of what is being taught. In other words, I understand him to say that the most efficient way to teach is by propaganda. This is perhaps not a conclusion that will be acceptable to all, but the exercise of reading how this conclusion is reached will produce many valuable new perspectives.

Despite the overall interest of the book, a vital spark which one feels should be there is missing. On reflexion, I felt that the missing aspect was information about what the educationists gained from the lectures, and how much relevance they felt that the lectures had for their task of education; too many of the contributors' remarks call for an immediate answer. And I wished that my brief had been to review the original live lectures rather than the subsequent publication.

C. VAN-TOLLER

Development and Evolution of Behavior. (Essays in memory of T. C. Schneirla.) Edited by LESTER R. ARONSON, ETHEL TOBACH, DANIEL S. LEHRMAN and JAY S. ROSENBLATT. San Francisco: Freeman. 1970. Pp. xviii + 656. £5.60.

Ethology: the Biology of Behavior. By IRENÄUS EIBL-EIBESFELDT. (Translated by ERICH KLINGHAMMER.) New York: Holt, Rinehart & Winston. 1970. Pp. xiv + 530. £5.00.

These two books are as contrasting as they are complementary. The first consists of 28 disparate contributions, dealing with many different facets of animal behaviour, but including also a few papers on human psychology. The animal studies comprise aspects of bee and ant behaviour, schooling in fishes, mating and maternal behaviour in the rat, as well as such general topics as prenatal behaviour, imprinting, orientation and communication. Many of the papers represent a characteristic approach to the study of animal behaviour of which Schneirla was the leading exponent. Contrary to what is sometimes said by its critics, this approach is not to be identified with the behaviourist tradition in psychology. It is a little inconvenient that as yet no single snappy label has been attached to the Schneirla-type thinking about comparative

psychology. The second book has an organization and unity about it almost suggestive of a textbook. It is, in fact, a comprehensive exposition of another tradition in the study of animal behaviour, that of ethology, whose founding fathers are Lorenz and Tinbergen. Very different though it is, this book, too, ends up with a section on human psychology, called the ethology of man.

Schneirla's work does not perhaps receive as much general recognition as it deserves. *Principles of Animal Psychology* by Maier and Schneirla, first published in 1935, is still today a great systematic reference work. Schneirla's studies of the behaviour of ants are of a pioneering character. Above all, Schneirla gave in his theoretical writings a certain distinctive flavour to American comparative psychology. He was highly critical of such key ethological concepts as 'releasers' and 'fixed action patterns', believing them to be misleading and experimentally unfruitful. As it turned out, these and similar concepts, crude though they may be, have proved most useful in initiating a very great deal of valuable research. Schneirla was equally critical of Skinner's brand of behaviourism. But in this case, too, operant-behaviourist approach has led to important results, with considerable theoretical and practical implications. Perhaps the most important positive influence of Schneirla's thinking has been his emphasis on the value of the developmental approach to the study of animal behaviour. He regarded it as more helpful than cross-species comparisons. To gain a genuine understanding of behaviour – Schneirla taught – one should study its ontogeny. Many today follow this precept.

A leading exponent of this view is Lehrman, whose essay returns to the theme of his earlier critique of ethology, published in 1953 in the *Quarterly Review of Biology*. The present essay is in the main a further critique of Lorenz's views, as expounded in the latter's book *Evolution and Modification of Behaviour*, which appeared in 1965. It is remarkable how much closer than at one time are Lorenz's and Lehrman's positions with regard to the hereditary and environmental influences in the ontogeny of behaviour. And yet, despite the changed stances of these two, and despite everything that has been written on the subject of the controversy by Hinde and others, the gap between the protagonists still appears to be considerable. Perhaps this is a case of failure of communication among scientists. Whatever may be behind this clash of views, it is nicely illustrated by the contrast in discourse between the writings of many of the contributors to the volume in Schneirla's memory and the book by Eibl-Eibesfeldt.

This last-mentioned work is impressive in its wide coverage, aided by many fine figures and plates. Dr Klinghammer, himself a well-known research worker in animal behaviour, has succeeded in making the book read not at all like a translation. The aim of the volume is ambitious in that it sets out to treat behaviour, both animal and human, strictly in the framework of ethology. As such, it often seems to sweep aside all modes of interpreting animal behaviour other than those within the ethological mould. Furthermore, the author accepts many sorts of evidence too readily and uncritically; this becomes especially so in relation to human behaviour, where diverse reports, suggestions and opinions are all treated as grist to the ethological mill. But one must hasten to add that, despite these misgivings, the present reviewer regards Eibl-Eibesfeldt's book as a most valuable survey and analysis of animal behaviour studies.

Too much is sometimes made of the controversy between the comparative-psychological and ethological camps. The pugnacity of the adversaries is partly to blame, but also semantic and conceptual difficulties. Additionally, there is a difference in the focus of attention; ethologists are particularly concerned with phylogenetic evolution of behaviour patterns; others are increasingly interested in the ontogenetic development of behaviour. This makes the two approaches essentially complementary. In any case, all empirical knowledge is common property, and whether it is viewed within one or another framework matters perhaps relatively little. That this knowledge may be plausibly presented in various ways is probably all to the good; and the lack of uniformity in approach is likely to lead in the end to a better-documented and more sure-footed science of behaviour.

W. SLUCKIN

The Social Behaviour of Animals. By S. J. DIMOND. London: Batsford. 1970. Pp. 256. £2.50.

Much of the present-day interest of biological scientists and non-specialists in the social behaviour of animals is with reference to the study of wild animals in their natural surroundings

and in the behaviour of wild animals in captivity. Moreover, interdisciplinary effort is now being concentrated upon the collection of reliable field data, and in extending the range and application of experimental techniques to problems suggested by naturalistic observations.

Dr Dimond considers the apparently inherent methodological and theoretical limitations of naturalistic observation. With reference to Goodall's (1968) report on 'The Behaviour of Free-living Chimpanzees in the Gombe Stream Reserve', for example, he writes: 'First, the investigator introduces uncertainty into the situation, in that his presence can influence the group structure and the nature of the behaviour he observes...Secondly, in most cases naturalistic observations do not permit a final analysis of causal relations... investigations are now needed which penetrate behind first-order observation into the causal structure of the communication system and the chain of command within the group' (pp. 165–6).

His own emphasis in this field is primarily concerned with the experimental manipulation of social behaviour and with special reference to the application of operant conditioning techniques. Hence, most of the experimental work which he cites relates to social behaviour in laboratory domesticants such as rats, mice and rhesus monkeys. He gives examples in which operant methods provide convenient ways of examining the effects of early experiences in the behaviour of infant animals, indexes of sexual and maternal behaviour, aggressive interactions and in investigating the processes and consequences of social learning in different species. In all these situations (and a variety of others, for which specific experimental examples are described) 'the extent of lever pressing provides an index of the value to the animal of the social stimulus' (p. 11). In his view, an integrated operant behavioural science has applications which extend far beyond that of operant conditioning techniques as convenient measures of learning in individual animals; it can make a substantial contribution to the development of behavioural science in general, including studies of social behaviour.

The main emphasis of the book does not, however, preclude the inclusion of references to naturally occurring social behaviours in a variety of species. For example, in the chapter entitled 'Group Behaviour' Dimond gives a brief account of social organizations in the wild, which he divides into five main categories: the reproductive pair, the family group, short-term and long-term larger groupings, and societies which are composed of 'many or all of these different types of group' (p. 157). Again, in his chapter on 'Sexual Behaviour', he mentions a variety of biological functions of sexual behaviours, the importance of hormonal control in sexual motivation, in addition to a number of examples to demonstrate that environmental manipulations can modify patterns of sexual behaviour in different species.

In brief, the book contains ten clearly written introductory chapters which indicate, and briefly discuss, examples of experimental work in areas of interest which include parental, sexual and aggressive behaviour, social learning, and relationships of social rank, including dominance, territoriality and leadership.

The familiar kinds of generalization, from animal studies to the social behaviour of man, are also discussed. In this connexion three main points are specifically important. First, there is the highly appropriate, if non-lucrative caution that human behaviour (as also the behaviour of any species) must be studied in its own right and on its own terms, rather than by a series of extrapolations from a conglomeration of findings from species in widely different animal groups. Secondly, that the application of experimental method, and of operant techniques in particular, is highly relevant to the study of social behaviour in any species, including man. Thirdly, that it is always important to study the *development* of human and infrahuman social behaviour, and especially so in view of the severe difficulties at present inherent in the 'learned *v.* innate' dichotomy.

It is always pleasant to read of different approaches to an area of inquiry which is still young enough to make it unwise to reject any methodologically and theoretically reasonable methods of study. On the other hand, it would also be unwise at this stage to run away with the idea that relative simplicity, objective experimental control, and the automatic quantification of data should take *a priori* precedence over other approaches in the study of intricately balanced and highly complex intra- and interspecific interactions. This point is well illustrated by the curiously oversimplified view of 'dominance' relationships which can be implied by presenting material from albeit convenient experimental techniques, which examine single pairs of animals in physically restricting environments.

O erant conditioning techniques undoubtedly have much greater potential value in the

study of animal behaviour than is commonly recognized, and clearly have the advantage of linking some attributes of individual motivation to specific environmental stimuli. All present approaches to the study of animal social behaviour, however, have their own difficulties. In the last analysis their efficacy obviously depends most importantly on the nature of the questions we wish to ask.

Finally, as for any technical book which is intended for lay readers and non specialist students, the inclusion of some illustrative material and a short glossary of terms (such as 'limbic structures') would have been useful.

HILARY O. BOX

Social Groups of Monkeys, Apes and Men. By MICHAEL CHANCE and CLIFFORD JOLLY. London: Cape. 1970. Pp. 223. £2.75.

This book provides a useful addition to the literature for the student who, wishing to gain an insight into the present-day status of research into primate behaviour, seeks a clear and lucid introduction to the subject.

One chapter sets out in a very readable fashion the different physical features of the primate groups. The rest of the book is devoted to description and discussion of the numerous field studies of primate behaviour undertaken in recent years. These are discussed in relation to their principal findings: behaviour of the young, movement within society, and the social structure of the group. Finally, generalizations are made about the nature of social groupings, and on this basis inferences are drawn about the nature of the social organization of man.

The authors show that the sequence of behaviour is important and they attempt to go beyond the mere cataloguing of the behaviour of individuals to discuss something of behaviour related to the social context and to answer questions about the nature of social control. This is a difficult task because many biases can enter into the original observations. Experimental reports may be unsuitable for this type of inquiry. Poor conditions of observation may limit the collection of data. The observer himself may drive the individuals from accustomed habitats and disturb the existing system. The observer himself may by his presence trigger much of the behaviour of threat, aggression and social alignment which he observes. There is also a danger, which seems already evident, that the undesirable features of some sociological surveys undertaken of man may be perpetuated in studying the social relationships of monkeys. However difficult the task, the challenge is an important one and we have already seen that ethological methods can make a significant contribution to the study of human behaviour.

The authors readily take up the challenge and their view is that non-human primate society can be categorized as centripetal or acentric organizations based on threat, aggression, and the dominance hierarchy. The controversial thesis is advanced, based on observations of the behaviour of rodents, that the threatened animal in an aggressive encounter heads towards the source of the attack – the more dominant animal. This, it is suggested, is the basis of centripetal organization. In respect of rodents the observations on which this argument rests are atypical and in the case of primates it is not clear that the suggestion is widely substantiated either. Whilst it was formerly accepted that dominance ranking was a pervasive feature of primate relations and the basis of group structure, today it is by no means as clear, and we have to ask if other social and familial aspects of behaviour are not equally or even more important. If there are difficulties in reaching agreement upon the nature of social organization amongst infra-human primates, then how much more difficult it is to agree upon the application of the principles to man. Those who remain unimpressed by the assertion of Lionel Tiger that the development of bonds between males represents a key to human society are not likely to be impressed with the claim that rank ordering and centripetal grouping could provide an innate component in the associations which develop between men in groups in our own complex society.

STUART J. DIMOND

Social Behaviour in Birds and Mammals. Essays on the Social Ethology of Animals and Man. Edited by JOHN HURRELL CROOK. London and New York: Academic Press. 1970. Pp. xi + 492. £7.00.

While ethology's classic textbook *The Study of Instinct* has found several modern and worthy successors, Tinbergen's other and equally important book *The Social Behaviour of Animals* has not yet been replaced. Anthologies of papers published in journals or presented in symposia are therefore welcome substitutes, even though, in the reviewer's opinion, such collections rarely reach the standard of a good textbook. This book is composed of written versions of talks given at a seminar in the Department of Psychology at Bristol. It is dedicated to the memory of Professor K. R. L. Hall a distinguished contributor to the field of social ethology, and includes a brief biography and a list of his publications. The editor's introduction provides a short historical account of the ethological and non-ethological antecedents of modern studies of the social behaviour of animals with a warning regarding the currently fashionable and facile translation of findings on animals into explanations of human social behaviour. Compared, however, with his recent and excellent review in *Animal Behaviour* **18**, 197–209 covering similar ground, the introduction is disappointing. Veiled references to system analytical approaches to social behaviour, for example, are not useful to the reader unless it is explained what they involve.

Goss-Custard presents the results of the study of dispersion of wintering wading birds, particularly in relation to the distribution of their prey. For inclusion in a book of this nature the paper is too specialized, including as it does quite extensive statistical tables. A wider ranging review of the literature on the dispersion of birds generally would have been more satisfactory for the reader. Simmons describes and compares the behaviour of a tropical gannet and of the crested grebe in terms of their adaptation to the environmental niches these species occupy. Only a small part of the conclusions is relevant to the main theme of the book, and then their foundations are restricted, without a review of work done along the same lines in a number of other species. Aldrich-Blake discusses the difficulties inherent in studies of free-ranging, forest-inhabiting monkeys and draws particular attention to the biases these introduce into studies of short duration. It is concluded, taking this into account, that the earlier, assumed, clear-cut correlations between the social structure and the forest are more complex than previously thought. The paper is healthy in its criticism, even if disheartening. It leaves one wondering how many of the hard and fast statements one sees in the primate literature are more due to pressures of publishing than to well-founded knowledge. Crook, in an excellent and wide-ranging review of the social structure and ecology of primates, draws particular attention to the intraspecific variability of social organization. Social traditions are considered to be responsible in part for this diversity. Hence cultural evolution may, as it were, be superimposed on the genetic evolution as a determinant of social behaviour in primates, much as it might have been during the development of human society.

Archer, in a well-presented contribution, reviews the effects of increased population density on the behaviour of mammals where endocrinological modifications related to the stress syndrome have multiple consequences on all spheres of behaviour. In view of the widespread presence of this mechanism among mammals it seems not too tentative to suppose that they also affect man. The chapter by Crook and Butterfield on 'The gender role in the social system of quelea' is again a specialized report on some observations and experiments on a species of finch. Although Crook makes the interesting suggestion that the application of the role theories to animal behaviour might be fruitful, the use of the word 'role' in this context seems to be a long way from that current in the social psychology literature.

The contribution by Butterfield is a somewhat meagre review of the evolution of pair bonds in birds and some inconclusive experiments regarding the maintenance of pair bonds in a finch. Vine's chapter is an extensive, useful review of communication by facial visual signals. Beginning with a brief overview of this type of signalling in animals, he examines critically its occurrence and significance in man. He concludes that it is very important for the maintenance of social behaviour in humans but that its causation and functions are multiple and that no single theory can explain all aspects of it. Kear, reviewing the parental behaviour of a large number of waterfowl studied at the Slimbridge waterfowl collection, neatly exemplifies the power of comparative ethological studies for the elucidation of the probable evolutionary history of

[illegible] The argument is lucid. The contribution by [illegible] presents some rather [illegible] experiments on juvenile birds. I can only recognize some as relevant to [illegible] and in any case they require an effortful immersion in extensive technical [illegible]. [illegible] reviews some of the theories which have been proposed to explain imprinting and the early experience which crucially affects the socialization of a number of animal species. Based on results of embryonic stimulation experiments, he proposes one of his own, but the [illegible] is very condensed, making it difficult to assess its value. The book thus contains a number of fine chapters, but it also includes some specialized contributions of restricted scope and covers the area of its title rather unevenly, leaving many topics untouched and even unmentioned. Prospective buyers should also consider the rather staggering price.

JUAN D. DELIUS

Traumatic Aphasia: Its Syndromes, Psychology and Treatment. By A. R. LURIA. (Translated by DOUGLAS BOWDEN.) The Hague. Paris. Mouton. 1970. Pp. 479. 96 Dutch guilders.

This book was originally published in Moscow in 1947 and a somewhat crude English translation has been circulating in typescript for some years. The text has, however, been considerably revised and the present translation is altogether more professional than its unofficial predecessor. The bibliography has also been brought up to date.

As students of what Dr MacDonald Critchley in his foreword dubs 'Aphasiology' will know, Professor Luria's work is characterized by perceptiveness and vigour, and by truly international scholarship. His point of view derives in about equal proportions from Pavlovian 'neurodynamics', classical German neurology and the English evolutionary tradition as created principally by Hughlings Jackson. It is difficult to believe that any other student of aphasia now living could combine so happily these so very different patterns of intellectual outlook.

Basically, this book embodies a study of language disorders due to penetrating wounds of the brain undertaken by the author in the course of the last war. In successive chapters he considers the nature and early course of traumatic aphasia, its classification and 'topical syndromes', methods of neuropsychological investigation (mostly qualitative but some highly novel), disorders of reading, writing and calculation, and the rehabilitation of aphasic patients.

Theoretical issues obtrude throughout the work and there is no attempt to shirk the exceptional difficulties involved. By and large, Professor Luria maintains a sensible compromise between the claims of the 'localizers', who assert that the various components of language are discretely localized in the cerebral cortex, and those of the 'holists', who see language as an essentially global cerebral function. He postulates a number of 'systems' possessing a considerable measure of functional autonomy that serve to organize and co-ordinate the basic sensory, motor and intellectual elements involved in language. These systems possess a very real degree of cortical localization and may consequently break down in a relatively circumscribed way as a result of focal injury to the brain. Although their detailed nature is left vague, the general approach has at this stage much to commend it. At the same time, it is clear that Professor Luria has little use for those who, like Marie and Head, see in aphasia evidence for a basic intellectual, as opposed to a linguistic, deficit.

In spite of its exceptionally wide coverage, the bibliography gives evidence of somewhat hasty preparation. There are also some errors. For example, W. H. Broadbent, the first to describe pure alexia in 1872, should not be confused with D. E. Broadbent, our leading contemporary British experimental psychologist. One may also call attention to a touching reference to a paper published in 'The Journal of Nervous and Mental Disasters'.

Professor Luria is the foremost representative of what has come to be called 'neuropsychology', i.e. the use of findings from experimental and clinical neurology to establish general principles of behaviour. Although often dismissed as yet another form of reductionism, modern 'neuropsychology' differs from its predecessors in eschewing reflexology and in accepting the brain as the instrument of real human behaviour. The present book merits the attention not only of clinical psychologists and others directly concerned with the aphasic patient but of all psychologists who seek to base their understanding of thought and language on the realities of the brain.

O. L. ZANGWILL

Brain Damage and the Mind. By MOYRA WILLIAMS. Harmondsworth: Penguin Books. 1970. Pp. 174. 35p.

Dr Williams exhibits considerable courage in attempting to describe in 174 pages what Luria did in 470 pages. But it was a necessary and valuable task that she undertook, for there is no short work in English that collects the existing information on the effects of focal lesions of the brain on behaviour. And, at the present moment, our information about the role of the human brain in what might properly be called psychological, as distinct from simple behavioural processes can only be obtained from those unfortunate natural experiments on patients resulting from trauma or disease. The book is likely to achieve the status of a text, a brief, but essential, entry to the literature.

It is inevitable that an attempt of this kind should not be entirely successful. The material is complex, many observations are crude, having been made by neurologists, neurosurgeons and other experts who often lack psychological knowledge. They are not, of course, to be blamed for this, for our psychological knowledge is in many places so inadequate that ordinary language psychology best serves the neurologist.

The author recognizes the extensive material with which she deals to be difficult to present easily and systematically. To simplify her task she has chosen to consider, in separate chapters, disturbances in consciousness, memory, perception, motor activity, and intelligence and personality. She sees the categories to be conventional and that, with a given brain lesion, there is no necessary restriction of disturbance to one of them; a lesion may, for example, produce simultaneous disturbances of memory, perception, and speech, even though one of these functions may be more disturbed than the others. But there is value in ordering the material in psychological categories, no matter how crude. For to order it under brain regions, which is the alternative, might encourage the concept of strict localization of function: a pitfall the author implicitly avoids in her text. Nowhere does she attempt a strict one–one correspondence of function and brain locality.

Nevertheless, the book may give beginners the view that the neuropsychologist is concerned simply to correlate particular brain regions with particular psychological functions. Obviously, this kind of knowledge is of great value for clinical neurologists and psychologists who wish to localize lesions. But for those less concerned with clinical questions, and for whom the book is written, it might have been valuable to discuss explicitly the general problem of localization of functions. And, perhaps, the naive hydraulic brain model, which is little more than a description of some of the facts, could have been omitted and the space better used for this purpose.

Disordered functions occurring with brain damage are carefully and analytically described. Recent attempts to study these disturbances experimentally are well reviewed. Often these studies suggest analyses of function which may not have occurred in the first place to experimentalists working with intact persons, and, on the other hand, concepts taken from the experimental laboratory are shown frequently to aid analysis of a defect. Indeed, the value of the book for non-clinicians may be in such suggestions for new analyses (or a confirmation of old) rather than in the correlation of cerebral regions with function. And it may stimulate young workers to explore the possibilities of work in this field in order to gain better understanding of normal processes.

But there is also value in the book for clinicians. It exhibits clearly how some of the disturbances encountered by the clinical neurologist may be analysed with a precision not available to the psychologically ignorant, and that the power and extent of such analysis is increasing. Indeed, the recent contribution of experimental psychology to diagnostic neurology is demonstrated in this book to be considerable. And the collection of this material in an easily accessible form publicly establishes the method of analytical clinical neuropsychology developed largely by Zangwill and his students in the West, and Luria in the East. It is a more useful and sophisticated method than that of the early 'testers', who could, after the neurologist had already made his diagnosis, sometimes identify 'brain damage'. This is a caricature, but one well embedded in many neurologists' (and some clinical psychologists') thinking. It may be eliminated by this book.

The writing is lucid and undogmatic. There are some references to the experimental work on animals and on humans without brain lesions. These references tend to be superficial and are superfluous, for it is questionable how far studies of animals can illuminate complex human

verbal and symbolic processes with which the work is largely concerned, even in the chapter on motor action.

The book will be valuable for those beginning the study of its field, and most certainly for psychologists who intend to work in the neurological clinic. The bibliography, dated up to [illegible], is short but a sufficient introduction to the very extensive literature.

The word 'mind' in the title may irritate the hard nosed, but, to distort Wittgenstein, can anyone think of a better word to denote those observations of persons discussed by the author?

R. C. ALBINO

The Grassi Block Substitution Test for Measuring Organic Brain Pathology. By JOSEPH R. GRASSI. Springfield, Ill.: Thomas. 1970. 2nd edition. Pp. vii + 84. $6.50.

Clinical Psychological Assessment of the Human Figure Drawing. By MARK McELHANEY. Springfield, Ill.: Thomas. 1969. Pp. xi + 256. No price given.

Both of these books are beautifully produced; the same publisher is responsible for both and deserves high praise. But, regrettably, the same cannot be said for their contents.

In the first book Grassi describes his Block Substitution Test; it consists of a set of five semi-cubes and four individual cubes similar to those of the Kohs Block Test. The semi-cubes contain patterns on the top, sides and bottom. Instructions for administering the test and for interpreting the results obtained are given. Eleven illustrative cases are discussed (36 pp.). The standardization of the test is briefly described in nine pages.

The introduction, of 14 pages, includes a discussion of concrete and abstract behaviour, largely based upon the work of Goldstein and Scheerer. Zangwill's review of the Goldstein–Scheerer tests in the *Third Mental Measurements Yearbook* is partially quoted, but no mention is made of Payne's more recent and more critical review in the *Sixth Mental Measurements Yearbook*: Payne remarks that so-called abstract ability, as measured by the Goldstein–Scheerer tests, is difficult to distinguish from general intelligence. This is important because one feels that the same may be true of the Grassi test. As Grassi himself remarks (p. 16), the Goldstein–Scheerer Cube Test 'is the only other organic test employing cubes (Kohs)'.

Grassi quotes (pp. 66–67) some results derived from 30 lobotomy patients on the Wechsler-Bellevue Scale and his own test, which surprisingly suggest that in these patients at least there is a negative correlation between W–B IQ and results on the Grassi test. But he does not give any more general figures for the correlation between his test and intelligence. This omission is important because he states that 'Analysis of test results is based on three related factors – behaviour, test score and intellectual level' (p. 25).

Grassi's basic claim is, of course, that his test differentiates normals from organics. This claim is supported by his standardization figures. Moreover, a paper is quoted on the back flap of the cover which states that 'validity data on 276 patients show a clear separation of organics from normals and undeteriorated schizophrenics with little overlapping'. But neither the author nor the data of this paper are mentioned. Grassi quotes two independent studies in the text which seem to support his claim, but it is surprising that out of the 18 references given to Grassi's test in *Personality Tests and Reviews*, edited by Buros (1970), only two are mentioned by Grassi himself. There seems, however, to be no doubt that the Grassi test differentiates between *groups* of patients, but whether it 'depicts early and minimal signs of impairment' in *individuals*, as is claimed, is open to question.

The second book, by McElhaney, consists of a single-page introduction and a five-page chapter on the rationale of human figure drawing interpretation. The rest of his book is taken up with case histories (including diagnosis, W–B intelligence and the Human Figure Drawing Test results). McElhaney introduces his book by quoting an unknown philosopher as saying 'One picture is worth ten thousand words' and then he remarks that assuming this philosopher is correct then 'there is the equivalent of more than 1,070,000 words condensed into this volume of 251 pages'. The rest of the book is no less modest.

McElhaney claims that the following information may be obtained from his test: (1) estimate of intelligence; (2) self-concept: (*a*) masculine–feminine, (*b*) grandiosity, (*c*) inferiority, (*d*) seductiveness, (*e*) rebelliousness; (3) concept of others; (4) hostility; (5) reality testing; (6)

[illegible] shown in this [illegible] through the test will prove rewarding to the clinician. [illegible] to some extent

It is [illegible] that any aspect of a patient's behaviour [illegible] that the claims made in both these books are exaggerated and unscientific. [illegible]

The Measurement of Psychological States through the Content Analysis of Verbal Behaviour. By LOUIS A. GOTTSCHALK and GOLDINE C. GLESER. Berkeley and Los Angeles: University of California Press. 1969. Pp. 317. $12.50.

Manual of Instructions for Using the Gottschalk-Gleser Content Analysis Scales: Anxiety, Hostility, and Social Alienation-Personal Disorganization. By LOUIS A. GOTTSCHALK, CAROLYN N. WINGET and GOLDINE C. GLESER. Berkeley and Los Angeles: University of California Press. 1969. Pp. 176. $6.00.

These companion volumes outline the authors' extensive research into the assessment of psychological states through the content analysis of verbal behaviour. The first volume discusses the theoretical framework, studies of scale reliabilities, construct validation studies, normative studies and examples of the application of the content scales in psychotherapy, psychoanalysis, psychophysiology, psychopharmacology and many other areas. The three main scales are standardized rating systems designed to assess the strength of anxiety, hostility and 'social alienation or personal disorganization'. In addition to the detailed description of these scales, the authors are also concerned with the general strategies involved in making inferences about psychological variables from verbal content. The second volume provides comprehensive instructions for the use of content scales.

Both volumes are written in a clear and concise manner, and are very readable. Evaluative comments are, however, more effectively directed at the quality of the research which is described.

One point in its favour is that the research is concerned with a very relevant and important topic. Psychologists have concerned themselves, to a great extent, with demonstrations of the fallibility of assessment techniques such as the interview and clinical assessment, where inferences made from the content of communication are important. There seems to have been less stress on methods of reducing error. Gottschalk and Gleser have concerned themselves with this problem—the main aim of their work is to standardize and quantify the inferences made from verbal content. Their research work is quite extensive, and has been conducted in areas ranging from psychoanalysis to psychopharmacology. But there are a number of serious criticisms which can be made.

The research almost certainly suffers from over-dependence on psychodynamic, more specifically psychoanalytic, theory. Dynamic theory and clinical judgement, rather than empirical investigation, have determined the content categories in each scale, e.g. the anxiety scale includes sections for death anxiety, castration anxiety, separation anxiety, guilt anxiety, and so on. But these categories of anxiety, though regarded as being equivalent in terms of dynamic theory, are demonstrated by the data presented here to be independent of each other. In other words, the various scales are clearly not homogeneous. But the scoring of each scale involves the addition of scores from these separate categories to provide one total score. This means that the theoretical assumption of equivalence is regarded as more important than the empirical demonstration of category independence or scale heterogeneity. So that we have the problem common to the dynamic approach—theory being allowed to supersede fact.

There are at least two other major difficulties. One is that the research is spread over a wide area, but does not seem to have penetrated any one area very extensively. The latter strategy might have provided a better idea of the practical limitations of these scales. A second set of difficulties involves serious shortcomings of many of the specific projects which are described. Examples of these shortcomings are numerous. Reliability studies are based on very small numbers of judges. Normative studies of the scales are likewise based on small numbers and, as in the case of 'psychiatric patients', on heterogeneous groups. The comparison of these

[illegible] with psychiatric diagnosis is contaminated, since both scale and [illegible] using the same cues (e.g. items in the [illegible] scale [illegible] would tend to be [illegible] significant [illegible] of [illegible]. Thus it is hardly surprising that [illegible] with [illegible]. [illegible], the concept of 'social [illegible]' or '[illegible]' is one which is popular in the dynamic field, but not derived from empirical work [illegible] the [illegible] which differentiate normals from psychiatric patients. [illegible] point is that weights are attached to the various categories which are used in each scale, but these weights are determined initially by the judgement of the researchers, and [illegible] which is both meagre (being based on small numbers and not cross-validated [illegible] and many others which cannot be [illegible] here, are so [illegible] that the [illegible] made in a large number of the investigations might be valid, but cannot be said to be so on the basis of the data which are presented.

In summary, the aims of this research are to be welcomed. But the theoretical bases and its [illegible] much of the work, the wide but sparse coverage of the research, and the [illegible] of many of the investigations, tend to reduce its contribution to the study of content [illegible]. In spite of these reservations, these volumes merit some attention from anyone interested in the field.

DAVID GRIFFITHS

The Psychological Assessment of Mental and Physical Handicaps. Edited by PETER MITTLER. London: Methuen. 1970. Pp. xxviii + 857. £8.00.

In the minds of many psychiatrists and social workers, the psychologist often appears to be rather like the man who comes to read the gas meter; he does not appear to be an applied scientist helping the clinic to do a day's work.' This remark, taken from a chapter by M. B. Shapiro, highlights a common attitude towards the psychologist who specializes in behavioural abnormalities, and the conflicts which he often faces. In an excellent introductory chapter on principles of psychological assessment, H. Gwynne Jones argues that the task of the clinical psychologist should be: the formulation of clearly defined hypotheses derived from the pool of data available about an individual patient, leading to the selection from a multitude of standardized tests those which are most likely to be relevant in this particular case; the interpretation of the data thus obtained, with possible extension of the evaluation, in the light of reformulations of the hypotheses; and the communication of his findings to his colleagues in such a manner as to be both clear and relevant.

This approach to psychological assessment demands the discipline of the experimental scientist, coupled with a certain sensitivity of subjective judgement. The applied scientist differs from his academic colleagues in that he must be prepared to reach a conclusion from incomplete data. He must work in a scientific spirit, be sceptical, test assumptions and predictions whenever possible and seek reasonable alternative hypotheses to explain the facts. He must also accept the responsibility of forming an opinion on less than adequate evidence and be ready to act on that opinion. In addition, as this book shows, he must be in possession of a great deal of technical information, and will probably be wise to gain specialist knowledge of a limited number of abnormalities, in order to be effective.

Thirty-five British psychologists have contributed to the book, which presents an exceptionally comprehensive account of the theory and practice of assessment of adults and children with a wide variety of impairments, including the deaf, blind, geriatric, mentally subnormal, language-impaired, maladjusted, mentally ill and physically handicapped.

This is mercifully not a cook-book of tests, although there are good summaries of recent research on the purpose, validity and reliability of intelligence tests, personality assessment, methods used in vocational guidance, the interview and psychophysiological assessment techniques.

A significant part of the text is devoted to accounts by clinical and educational psychologists of how they apply developing theories and on how new theories and techniques are developed. Thus, for example, Elizabeth Warrington discusses the cerebral organization of psychological functions with reference to systematic and experimental investigations of behaviours related to neurological deficits. The purpose of the researches was theoretical, but the test procedures evolved may be directly applicable in a clinical situation. Moya Tyson shows how, in the

development of remedial educational programmes, it may not be possible to [illegible] the exact area of disability until the programme is under way. Thus the teaching itself may carry the burden of diagnosis.

New procedures in assessment include, for example, the repertory grid method and the [illegible] interpretative than discussed by Bannister, [illegible] methods of automated testing systems which in some circumstances may prove to be a useful way of obtaining important information on psychological processes without involving the psychologist in time-consuming and often tedious routine procedures (Gedye and Miller).

Although individual authors show considerable differences in their analyses of problems, an important feature of the book lies in the presentation of carefully documented evidence on the redefinition of the clinical and educational psychologists' role. There are [illegible] who advocate a critical reappraisal of what has, on occasion, been referred to as the 'pretty games and puzzles' approach. As the editor puts it, psychologists in the past have tended to confine themselves to assessment, leaving others to incorporate or ignore their findings in whatever was done for the patient or client. The emphasis of the future will almost certainly involve the psychologist in using his own findings and theoretical knowledge as a basis for helping the individual patient. For this, methodologies concerned with the experimental investigation of single cases, advocated by Shapiro and others, need further development.

Readers will find this book a valuable source of recent research data and new theoretical orientations which should provide a most useful text for all those concerned with the psychological assessment of handicapped individuals, and for their colleagues who wish to understand what they are doing.

ANN M. CLARKE

The Paranoid. By DAVID W. SWANSON, PHILIP J. BOHNERT and JACKSON A. SMITH. Boston: Little, Brown; London: Churchill. 1970. Pp. xii + 523. £6.25.

The term 'paranoid' is generally used in a restrictive way to exclude fully developed paranoia and to limit its meaning to minor personality disturbances and conditions which are nevertheless of the same psychological order. There is, of course, no reason why such a limitation should be made, since psychological rather than strictly psychiatric studies have convinced us of the continuity between normality, slight and partial conditions of a schizoid, manic-depressive or paranoid kind, and the fully developed disorders of schizophrenia, manic-depressive insanity and paranoia which come under the psychiatrist's care. In this book the term 'paranoid' is used in its inclusive sense to cover all degrees and kinds of deviation and disorder of a certain and familiar kind, and which the authors take a great deal of trouble to define and explain very clearly.

The book is divided into four parts: (1) a general survey of the problems, in two chapters, dealing with the concept of the 'paranoid', and its history; (2) a review in eight chapters of the clinical aspects of the paranoid; (3) four chapters on theoretical aspects; (4) a survey in seven chapters of the influence of the paranoid in the family, social life and history.

The organization of the book gives an extremely clear and useful presentation of the whole subject. The problems of definition and the history of the concept of paranoia are very well handled. The discussions of clinical aspects are excellent, apart from one rather worrying point. This is that every disease or disturbance which has, or may have, paranoid components or ingredients or associated symptoms is discussed in great and illuminating detail, but often in such a way that the reader might have the impression at times that certain physical diseases, accidents, events, etc., were in themselves causes of paranoia or paranoid symptoms. We must remember that it is the impact of a physical illness or event on a certain kind of personality rather than another kind which produces paranoid results, not the physical illness or event in itself. The psychological testing of the paranoid patient is usefully dealt with in a chapter by Frank A. Dinello. Somatic and psychotherapeutic treatments are well described, and their limitations and advantages clearly presented.

The discussion of theoretical considerations, like the previous discussion of precipitating causes in the form of physical illnesses and events, again often gives the curious impression that physiological and sociological factors might in themselves be the essential rather than the precipitating causes of paranoid effects, and one wonders sometimes whether the authors really believe it.

[illegible] of the most interesting and useful parts of the book is the study of the influence and [illegible] of the paranoid within family life, in citizenship, government, politics, law, social reform, [illegible] life, [illegible], medicine, education, and its important imprint on history. The paranoid [illegible] with us at every turn, has been and always will be, as far as we can judge, and a careful and objective survey of paranoid influences, such as part IV of this book, deserves wide attention and thought. The value of this does not lie so much in the interesting discovery that such and such a well known figure was almost certainly paranoid, but in order to help us to perceive the nature of certain influences now at work, and to understand why and how they [illegible] to meet with such a surprising amount of support and approval, and to think how they can be contained by society.

R. W. PICKFORD

Behaviour Studies in Psychiatry. Edited by S. J. HUTT and CORINNE HUTT. Oxford and New York: Pergamon Press. 1970. Pp. ix + 223. £2.50.

This important book is based on the proceedings of a symposium held at the Human Development Research Unit, Oxford, in 1968. It is possibly the first complete work to apply the ethologist's methods of behaviour study to psychiatry.

In the opening chapter S. J. Hutt forcefully criticizes the many experiments in which behavioural measures are far less precise than they might be, particularly where they are to be correlated with physiological data, as in drug studies. Ethologists always commence by observing animals in their natural surroundings. Behavioural researchers in psychiatry could well adopt the approach and techniques of ethology when attempting to develop precise methods for analysing a patient's behaviour in his natural habitat. The ethologist, we are told, asks four key questions about behaviour; these concern its causation, present function, evolution in the species, and its development in the life-history of the individual.

It should be pointed out at this stage that, although the title does not give this impression, much of the book (six of the 11 chapters) reports work in the field of childhood autism. As the symptomatology of this condition is not exclusive, many features being present also in adult psychosis, for example, readers are likely to find much that is relevant to their own sphere.

A well-known concept of ethology is that of dominance hierarchy. Esser, applying this to the behaviour of male schizophrenic patients, produces an interactional hierarchy based on the amount of verbal and non-verbal behaviour on the ward. Examining territoriality and dominance behaviour he shows that the outcome of aggressive encounters is not as clearly related to position in the hierarchy of the patients as it is in animals, largely due to the role-playing ability of the former. This paper is accompanied by very full details of the method of observation, coding, and recording of data. Although this has resulted in some charts containing microscopic entries, the effort to share technology is appreciated. The latter may be even more true of Grant's contribution, describing in detail the construction and use of an 'ethogram'. Observations of the non-verbal behaviour during interview of psychotic and non-psychotic patients, and normal controls, are tabulated in the form of a contingency table for each group. This is a matrix giving the probabilities of each act preceding or following every other act. Four basic groupings of elements, essentially similar for each group, are: flight, assertion, contact, and relaxation. In the following chapter Currie and Brannigan make intelligent use of these findings in the study of the repertoire of an autistic girl. In view of the reawakening interest in single-case research it is a pity that this is only a pilot study. They report some success, however, in the application of operant reinforcement to certain missing links in the behavioural chain.

Castell shows some important differences between normal, brain-damaged, and autistic children on measures of mobility and attention in a familiar setting. One of these concerns the shortness of visual fixations on an adult by the autists (2–3 sec./min.). Its significance is further studied in 11 autistic children (using cine film) by Hutt and Ounsted, who relate it to social encounters in the normal environment. Coss reports in full four excellent studies, using normal subjects, which measure CNS arousal resulting from exposure to various configurations of 'eye-spots'. Two horizontal circles, schematically representing staring eyes, evoked significantly higher responses. The discussion at the end of this paper, as generally throughout the book, is germane and informative.

Three chapters are linked, in that all are concerned to some extent with the origin and function of stereotyped behaviour in autism. Greenbaum observes this, together with five other categories of behaviour, in six autistic children, attempting to detect any periodicity. Hutt and Hutt, in their valuable contribution, relate stereotypy to level of arousal, presenting extracts from concurrent film and EEG records. Stroh and Buick's attempt to reduce stereotypy by sensory isolation, whilst admittedly a pilot study, offers no quantification, controls or validation.

Containing more than 300 references, over 75 per cent of which fall within the last decade, the book should prove stimulating to all those currently concerned with psychopathology, whether they are students, research workers or clinicians. E. WHELAN

The Roots of Individuality: Normal Patterns of Development in Infancy. By S. K. ESCALONA. London: Tavistock Publications. 1968. Pp. 547. £5.00.

This book combines two topics of current interest. It describes the behaviour of two groups of infants in the first eight months of life, and compares them on 'activity level' behaviour connected with arousal.

The total sample of 32 infants was made up of the two most active and the two least active babies at each month of age between four and 32 weeks, selected from the records of 128 normal infants. These had been the subjects of a cross-sectional study conducted during a number of years. The infants had been assessed on developmental tests, physical growth and pediatric measures, and on kymograph recordings and ratings of activity before and after feeding. Detailed observations of each infant and mother during a four-hour period were made by independent observers. Films, an attitude questionnaire by the mother, and records from a home visit were also available. The emphasis was throughout on the minutiae of behaviour and on the contexts in which they occurred. 'Activity level' referred to the amount and vigour of bodily motion typically shown by an infant in a variety of situations, assessed by the pooled ratings of three independent observers.

In part II of the book the active and inactive babies are compared. Reactions to hunger and fatigue, to stimulation in different sense-modalities, and to inanimate objects and the frequency and manner in which these had been presented; responsiveness to social stimuli, spontaneous behaviour when awake and contented, and the extent to which infants stimulated themselves by sucking or touching are discussed. Accounts are given separately for the 4- to 8-, 16- to 20-, and 24- to 32-week-olds, supplemented by graphs and frequency tabulations in terms of age and active versus inactive status.

One of the aims was to relate individual differences to recurrent patterns of experience in the infants' daily lives. Part III describes each individual infant. The frequency of bodily arousal due to external stimulation and to bodily discomforts respectively, and the relative importance of near and distance receptors in the way the infant was stimulated by his mother are examples of experience patterns. These are discussed and tabulated in relation to the mother's competence, style of approach, and frequency of contacts with her infant, and in relation to the infant's developmental achievement scores and general irritability and excitability.

Dr Escalona's study provides a massive amount of descriptive material. We need not therefore quarrel with the fact that few statistical evaluations of the data are attempted, or, where they are, show relatively low levels of significance, and mainly between measures directly involving the criteria on which the two groups of subjects were selected. Dr Escalona may be justified, too, in considering that the imbalance of boys to girls, or of the distribution of socio-economic levels, is not important in a study which concentrates on description of the minutiae of individual behaviour and contexts. Dr Escalona uses the tabulated material primarily to summarize the frequency with which given behaviours occurred, and presents her conclusions as hypotheses which need to be tested further. The cross-sectional approach is to be regretted as there was a difference between some of the age-groups on the activity level ratings. It could be wished that at least some longitudinal data had been available.

The nature and amount of material to be presented must have made this an exceedingly difficult book to write. It is also somewhat cumbersome to read, especially the discursive initial chapters. An appendix, giving a succinct account of methods of assessment and ratings, would have been more useful than the lengthy and somewhat scattered discussions. The layout requires an inordinate amount of cross-reference. Dr Escalona leans towards psychoanalytic

ego theory and Lewinian formulations. Potential readers will differ in whether they regard these as helpful shorthand or unnecessarily difficult jargon. But Dr Escalona is not preaching a doctrine. Her point that behavioural 'states', individual differences in arousal level and the kind and amount of stimulation to which infants are habitually exposed should be treated as separate parameters is one which experimenters can ill afford to disregard. The material Dr Escalona presents in this book repays study.

SUSANNA MILLAR

Parental Attitudes towards Exceptional Children. By HAROLD D. LOVE. Springfield, Ill.: Thomas. 1970. Pp. ix + 167. No price given.

Dr Love has attempted to fill what he sees as a gap in the literature on exceptional children the absence of a 'satisfactory book which dealt with parental attitudes towards all areas of exceptional children'. He writes for an American audience and much of what he has to say regarding provision for exceptional children is relevant mainly to educational systems and social philosophies operating in the U.S.A. He has cast his net wide and includes in his list of 'exceptionalities' all physical and mental handicaps, sensory defects, 'special health problems', 'emotional and social disturbance' and giftedness.

Dr Love rightly emphasizes the importance of parents in the lives of all children but too often falls into the trap of repeating uncritically the gloomy list of pathological responses to handicapped children which their parents are so frequently alleged to develop. When he attempts more general discussion of the reasons for parental responses such as guilt, rejection, over-protection and the rest, his thinking and his prose become extremely muddled.

Research findings on parental attitudes are, admittedly, difficult material to handle, based as they are on very variable sampling and methodology. For these reasons a critical approach would have been welcome but again this book is disappointing. There is little attempt to weigh the relative merits of differing research approaches or to evaluate conflicting results. There is, however, a considerable amount of research (mainly American) listed and such references are useful.

The book is intended for 'educators, physicians, psychologists... and parents', but the discussion of IQ testing and the handling of psychological theory are both too superficial to interest the professional and too confused and jargon-bound to inform the layman. This is a pity, because Dr Love's intentions are sound and his work is redeemed from time to time by flashes of common sense, humanity and coherently presented information.

S. H. HEWETT

The Subject of Consciousness. By C. O. EVANS. London: Allen & Unwin; New York: Humanities Press. 1970. Pp. 240. £2.75.

Behavior and Conscious Experience: a Conceptual Analysis. By KENDON SMITH. Athens, Ohio: Ohio University Press. 1969. Pp. 139. $6.00.

According to some writers, there is a scientific revolution occurring within psychology; if so, these two books are probably near the forefront of the revolution. They are both concerned with understanding and dealing with consciousness, they are both written from within a general orthodox tradition (analytic philosophy and behaviourism, respectively), and they both find it necessary to make major leaps away from orthodoxy. In these respects the books are similar; in approach, content and merit, however, they are very different.

Professor Smith sets out to review modern behaviouristic practice, as revealed chiefly in textbooks, in order to find the true definitions of 'stimulus', 'response' and 'learning'. The chief fault he finds with all previous definitions is that they are not explicitly dependent upon conscious experience. This fault should be understood as one of form rather than of content; behaviourists have *really* been talking about consciousness all along, but they have 'repressed' the notion of consciousness, have exhibited a 'pathological hypersensitivity' to it, and have often revealed an 'explosive abreaction' to the content of their repressions (p. 128). The problem of definition can be solved, therefore, by making an explicit central reference to conscious experience instead of to such circumlocutions as 'adaptiveness'; in all other respects, basic S-R formulations can remain intact. Thus Smith defines a response as 'any liberation of chemical energy by a living object, as produced by conscious experience'. A stimulus is 'any

event, within or upon a living object, which has a material effect upon that object only insofar as it produces conscious experience'. The definition of conscious experience has to give meaning to all the preceding definitions, and also has to be appropriate for the 'sensations, perceptions, thoughts, images, dreams, wishes, and emotions' which Smith wishes to include in consciousness. The task is simplified a little, for he insists that all these contents of consciousness reduce to sensations and can be treated as equivalent. He defines consciousness (or conscious experience) as 'an internal event to which an arbitrary response can be attached directly by the process of following that response, when it occurs to that event, with a prespecified set of circumstances'. Now, if this definition is to mean anything at all, the 'prespecified set of circumstances' must be one which will have some effect upon the response; in the S-R formulation Smith employs, it must constitute a positive or negative reinforcement of the response. The definition thus comes down to 'any internal event which can be functionally related to a response through reinforcement', or even 'a sensation of a discriminative stimulus'. As a definition, this throws little light on the previous ones, and is plainly irrelevant to most of the range of conscious phenomena. If consciousness is to be dealt with conceptually, it will require far more than the simple fiddling with S-R formulations which, in the end, is all Professor Smith provides.

There are two basic approaches to problems of identity, awareness, cognition, etc. – the first-person (or *self*) approach and the third-person (or *persons*) approach. The first-person approach is based on immediate and non-inferential knowledge of the self, and the third-person approach is based on publicly verifiable, or at least publicly meaningful, propositions. In this century the first-person approach has been almost exclusively the domain of phenomenologists, while most other psychologists and philosophers have restricted themselves to the third-person approach. There are good reasons for this division of interest. The first-person approach is hard to reconcile with our notions of science as based on publicly observable events, and historical attempts to develop psychological and philosophical theories with the first-person approach have been mainly unsuccessful. Still, it is worth asking if these reasons are sufficient to warrant continuing the division. C. O. Evans contends that the two approaches are complementary, and that both are necessary in analytic philosophy. His book is an attempt to use the first-person approach in developing a theory of the relationship between consciousness, attention, and the self. In so doing, he hopes to throw new light on the philosophy of mind and, by implication, on the mind–body question.

To attempt such a theory is an ambitious undertaking, and requires a good deal of new analysis as well as reintegration of much old material. The author begins with a review and critique of 19th-century writings on attention and consciousness. The review leads into an excellent analysis of the structure and functions of attention. The author's theory of the self is dependent on the analysis of attention, and in quite an elegant manner identifies the self with those elements of consciousness standing in the background of attention. Dr Evans next discusses the philosophical and conceptual problems which his theory of self has some hope of resolving, and finally deals with the experience of continuity of the self. The book closes with some suggestions on the possible restructuring of relations between first- and third-person accounts.

An original philosophical analysis is an impressive task, and understandably some portions of this book are less adequate than others. The author himself is sometimes confused between the body as a source of experiences (first-person approach) and the body as an object (third-person approach). This confusion leads him, for instance, wrongly to reject disembodied existence as a logical possibility. More seriously, the chapter on the experience of continuity of the self is superficial and unconvincing. Nevertheless, this is an important book, and well repays careful reading. The theories are elegant in themselves and may have wide implications for semi-phenomenological research. The prefix 'semi-' is advisable; the greatest contribution of this book may be in aiding a reconciliation between the phenomenological and analytic traditions.

B. D. MACKENZIE

Personality: Theory, Assessment and Research. By LAWRENCE A. PERVIN. New York and London: Wiley. 1970. Pp. xiv + 632. £4.40.

The author has a strong sense of design. At the heart of the book are five chapters (5–9), each about a particular approach to personality: 'major theories in the field today. . . representative of the diversity of approaches that can be considered reasonable and useful' (p. 557). The theories

are those of Freud, Rogers, Kelly, Cattell and learning theory, representing respectively a psychodynamic, phenomenological, cognitive, trait (factor-analytical) and a variety of learning approaches. These central chapters use the same categories to facilitate comparison among the theories (a precedent set by Koch in his editorship of *Psychology, the Study of Science*). A pleasant touch is the provision of a photograph of each theorist (three for learning theory, namely Watson, Hull and Skinner because of their different influences on other learning theorists), with a concise, interesting account of their careers, their views of man and of science. Then come the categories: structure, process, growth and development (a format which suits some theories better than others). There follows an example of the way in which an individual would be studied in the particular framework, and a critical evaluation of the theory. Leading up to these central chapters are four concerned with (1) determinants of personality, (2) the nature of theory, (3) assessment (including principles and issues as well as methods) and (4) research. Each chapter ends with an 'overview' and the fourth adds a review of theory, assessment and research related together. After the central chapters comes a study of an individual ('The case of Jim Hersh') by means of the different tests which would be used by exponents of the five theories. The question asked by Pervin is whether the various approaches are disclosing different aspects of the person or whether the views are inconsistent. He concludes that 'in a certain sense, each represents a glimpse of the total complexity of human personality'. The final chapter reviews the content and conclusions of the book ('An Overview of Personality Theory, Assessment and Research'). Thus the design of the book is rather like a musical composition or dance where a motif is selected for separate treatment, recombined, recapitulated, with frequent reminders of the main theme and regular form ('Let us take stock of where we have been and of what lies ahead'; p. 204). However, this very determined carrying out of a clearly defined pattern serves well the aims of the author, including helping students 'to separate figure from ground...what is of major importance from what is of lesser importance'.

A query could be raised perhaps about the slightness of the mention of Festinger (p. 296) contrasted with the chapter on Kelly. No doubt it could be said that cognitive dissonance is not a personality theory, although as an approach to motivation and attitude change it is not irrelevant and it illustrates well the productive nature of theory-related research. Certainly a detailed account of Kelly's work is particularly useful now that his book is out of print and difficult to obtain.

Pervin considers that the best way to study human behaviour is from the standpoint of individual environment interaction (pp. 548–557). This does not prevent him from giving thorough expositions of very different approaches; for example, Cattell, and learning theory. He has an enviable capacity for objectivity.

This book is likely to be welcomed at all levels of study, being solid and also readable, with its clarity of design and style. Many useful minor theories (or those with a more limited range of convenience) are expounded in the chapters outside the central five. The chapter on determinants ranges from Sheldon to Kardiner (in keeping with an interactionist view). Principles for evaluating theory are stressed. There is a bibliography of 48 pages, including work of British psychologists (Bannister, Barbu, Eysenck, Vernon). It is a great advantage to have assessment methods, research and theory discussed in the same text, with equal competence. In his acknowledgements Pervin refers to 'the four years of work on this book'. One can but admire to what good purpose he spent these years.

RUTH BOWYER

Handbook of Personality Theory and Research. Edited by EDGAR F. BORGATTA and WILLIAM W. LAMBERT. Chicago: Rand McNally. 1968. Pp. 1232. £10.00.

What is, or could be, the subject-matter in that branch of psychology indicated by the title of this substantial book? There is some merit, as a description of what we have, in Albert Mehrabian's thesis that hypotheses investigated empirically in 'personality research' invariably derive from a general theory and that each theory is the elaboration of an arbitrary metaphor or analogy.

This leaves the field wide open to speculative builders and hence one cannot object strongly to any definition or approach unless it is manifestly absurd.

However, one might suggest that personality study, while not neglecting its inevitable debt to psychopathology and clinical psychology, could emphasize any individual differences of

normal subjects which can be identified and measured, assessed or, even speculatively, derived from data with only weak verification. 'Need for achievement', 'self-concepts', 'self actualization' or 'individuation' would seem to be obvious dimensions, but these cannot be appreciated without relating them to the theories of Murray, Rogers, Maslow and Jung from which they derive. 'Intelligence' and 'creativity', although equally belonging to psychometrics and cognitive psychology, come into the picture. But what else? Many of the grand theories, which obviously give the psychology of personality an existence, have not generated much empirical research or have put forward rather questionable dimensions for measurement or investigation.

This handbook of 1187 pages reveals this dilemma. The editors want to present research problems and the results of 'scientific' investigations and they succeed in assembling 29 authors who, between them, provide 24 review-article chapters. These contain a wealth of information and commentary and the work is well done: it is a tome worth having, worth studying and following up through the massive lists of references at the end of each chapter. However – and this is no fault – it ranges very widely in scope. Five chapters might occur in a handbook of clinical psychology (Concepts of Normality; Defense Mechanisms; Human Reactions to Stress; Syndromes of Deviation; Counselling and Psychotherapy); five others might constitute part of a work on social psychology (Childhood Socialization; Adolescent Socialization; Leadership; Affiliation and Social Comparison; Personality and Susceptibility to Social Influence). Throughout the book there are many discussions relevant to the student of developmental psychology.

The chapters on the physical bases for individual differences are certainly most useful. William Kessen in chapter 6 deals with comparative psychology and its relevance for problems of personality development: there are some new approaches here which are obviously worth development. David Lykken in chapter 7 has two discussions: (*a*) a review of physiological psychology in which independent variables are physiological manipulations (brain stimulation, drug injection, brain surgery), while the dependent variables are behavioural, i.e. test scores; (*b*) psychophysiological research in which behaviour is manipulated and physiological effects measured (e.g. stage-fright induced and heart-rate changes recorded). Part (*a*) consists of reportage of work on the reticular formation and the not unfamiliar data on cortical activation and behavioural arousal, etc. This does not seem to give much insight into individual differences of the kind students of personality have indicated. Part (*b*) is hopeful rather than rich in achievement but clearly new techniques for measuring basic individual differences of response are being evolved from which personality study will benefit. William R. Thompson's chapter on 'Development of the Biophysical Bases of Personality' is an authoritative review of the literature on 'developmental stratification'.

The four chapters devoted to theory are each detailed and thorough. Berlyne on 'Behaviouristic Personality Theory', Ruth Wylie on 'Self Theory', Edwin Thorn on 'Role Theories' and David Glass on 'Theories of Consistency' (Festinger's cognitive dissonance theory, Heider's balance theory, Osgood and Tannenbaum's congruity theory). In some ways this is an odd selection from among those currently available – but clearly even in a book of this size a limited selection had to be accepted.

One of the most interesting chapters is that by Richard Christie and Florence Geis, 'Some Consequences of Taking Machiavelli Seriously'. This concerns the problem of investigating those individuals who are 'effective in manipulating others'. Is there a type or group of types? How do they develop a set of traits which summate in the dominant need to 'manipulate'? What forms of behaviour does manipulation take? This would seem to be a distinctively 'personality' topic whatever the quality of the 'results' in the investigations discussed. It is interesting that Machiavelli as well as Adorno, Hoffer and Shils is among the theorists from whom the hypotheses in this chapter are derived.

Although J. C. Burnham begins the handbook with 76 pages of the history of personality study and Irvin Child has a scholarly chapter on 'Personality in Culture' there is no concluding chapter which attempts to analyse the state of this field – the parts which cohere and those which are fragmentary and peripheral, the flux and stability over the years in the 'problems' defined as 'basic', the possibilities emerging from an over-view, and so on. This is perhaps the only weakness in this excellent collection. There is no emerging picture of a field of related problems or topics. The principal need is for analysis and clarification of basic concepts and methodologies. What are the characteristics of a human being which constitute the basic dimensions along

which individual differences are to be measured or assessed? What sort of characteristics are these and how do they function in describing and explaining what people do? Are the metaphor-structured theories of personality from which hypotheses are worked up the right kind of theories to achieve this basic requirement of identifying individual differences of the sort psychology (as distinct from ethics or sociology or literature) requires? Is the psychology of personality a set of relatively distinct and coherently related empirical problems, or a rag-bag for categorizing odds and ends which do not conveniently fit in with clinical, psychopathological, social and developmental interests? The handbook does not resolve this sort of confusion but it contains plenty of data to interest most psychologists, and constitutes a useful work of reference.

R. THOMSON

Social Psychology. By BARRY E. COLLINS. Reading, Mass.: Addison-Wesley. 1970. Pp. x+389. £2.60.

This unusual textbook aims to examine certain related aspects of social psychology, namely social influence and conformity, the concept and measurement of attitudes, group processes and prejudice. Its coverage is selective, and as the author says in his preface, 'perhaps there was need for a textbook with more procedural descriptions and content than is available in most experimentally oriented texts... [and which] integrates methodological content with substance'. Given these self-imposed restraints, to what extent has this approach been successful?

Possibly the most striking aspect of this book is the descriptions of procedure – descriptions which are both careful and highly evocative of the reality of experimental situations. For instance, Collins quotes from Milgram's transcripts of conversations between the experimenter and subjects in his work on the effect of social influence on shock administration, thus providing an effective supplement to the data displayed.

The choice of areas is a promising one, since other textbooks often treat these topics in relative isolation, and an attempt to integrate them even partially is to be welcomed. The generally sophisticated level of discussion makes it all the more disconcerting – for an English reader at least – to find the interspersion of elementary definitions of 'variable', 'significance level' and other concepts commonly introduced in the first term of a psychology course. Moreover, the brevity and consequent over-simplification with which these are treated make them at worst misleading and at best inadequate for a thorough understanding of the rest of the text.

The series of propositions introduced throughout, as for example Proposition 6.4C, 'the presence of other people can constitute a distraction and lower productivity', is effective in drawing attention to various points, but would be more valuable if the findings quoted in support of each were assessed with regard to degree of validity and generality; in some cases the propositions cannot be fully justified or are ambiguous. In the case quoted, for instance, some reference to the literature on improved performance in group conditions could well have been referred to, since it is included in an earlier proposition.

The first of the four main sections, on social influences and conformity, ties in the findings in this latter area with those on attitude change and examines them in the light of other data on social influences. This, together with section II, on group processes, which provides a thorough coverage of task variables and motivational aspects and which includes some interesting material, is possibly the strongest aspect of the book. Sections III and IV, on attitudes and prejudice respectively, are less comprehensive, although this is understandable in view of Collins's intention of stringently selecting material. Section IV, contributed by R. D. Ashmore, attempts to show the applicability of social psychology to a 'real-life' problem, and effectively surveys various definitions and theories of prejudice as well as considering alternative methods used in studies with the aim of reducing it.

Social Psychology would probably be a stimulating textbook for more advanced students who had already some knowledge of social psychology; or it could be used in conjunction with a more basic textbook. The layout of the book and the use of tables and graphs are pleasing. It is unfortunate that the same cannot be said of the publisher's taste in cover designs.

ANGELA B. STEER

Managing for Accomplishment. Edited by BERNARD M. BASS, ROBERT COOPER and JOHN A. HAAS. Lexington, Mass.: Heath Lexington Books. 1970. Pp. x + 317. $12.50.

The papers collected in this volume were originally presented at a two-day symposium sponsored by the Office of Naval Research on factors making for effective management in complex organizations. Their publication is clearly intended to mark this event and to make the various contributions available to a wider public. Nevertheless, one cannot help feeling that they will be of greatest interest to those who were actually present at the symposium and who will be able to recall the context and substance of the resulting formal and informal discussions. As with all works of this kind, the differences between contributors in both style and content, makes for rather a patchy book with the theme of task orientated management as a slender thread running throughout the whole. There is no doubt that, as the editors state in their preface, understanding of the interrelations between the human side of enterprise and the task is essential to full comprehension of the role of the manager. What is presented here is a mixture of research papers, surveys of published work, and one or two more speculative articles, which will be of interest to academics working in this field. It is not a book for practising managers or those seeking guidance as to how to organize for task accomplishment. As a collection it records the thoughts of a group of research workers and social scientists at the beginning of the 1970s on some of the perennial problems of leadership and managerial behaviour and, in this respect, it will serve a useful historical purpose when it is viewed retrospectively in the later decades of this century.

Following a brief introduction by Bernard Bass, the papers are grouped under four headings: Individual Differences in Concern for Task Accomplishment, Stimulating Concern for Task Accomplishment, Organizing for Task Accomplishment, and Transnational Implications. Under the first heading a variety of research is reported which seeks to explore the concept and consequences of task-orientation in a work leader. This approach to some of the familiar problems of management seems to have potentialities, but, as Robert Cooper, the only British contributor to this symposium, points out, there is still confusion as to precisely what is meant by 'task-orientation'. For the concept to be of utility some consensus of definition is required – for example, between those who see it as essentially a production-centred activity and those for whom it denotes an individual's job competence and motivation.

In the second section the reviewer found a speculative paper by Karl Weick on 'Retrospect in Tasks' one of the most stimulating in the book. In it he argues that the meaning we attach to our actions is a retrospective phenomenon – that is, that man knows what he has done only after he does it – and discusses the implications of this idea for theory and research. One would have liked to see more papers of this kind in the book or at least one or two more which attempt to explore and test Weick's assertions. A research paper by Rosen in the same section is another which stands out from some of the more mundane contributions.

Part 3 opens with a survey of the literature on the relationship of task and human conditions to effective organizational structure by Friedlander, which is another useful contribution. Mouton and Blake also have a paper in this section and another in part 4, dealing with transnational organizational development. This seeks to show, not altogether convincingly, that across countries where data have been obtained, the 9,9 way of managing human resources in managerial grid terms is 'universally' endorsed as the most preferred and most effective approach. One cannot help feeling that if this were so, international companies would find their managerial problems much easier to deal with than appears to be the case. In fact, interesting and important as these transnational implications are, it is this concluding section of the volume which highlights the limitations of our knowledge. Perhaps it will be this aspect which will show the greatest development in this decade. At present we are not in the position to do more than put out tentative feelers and indicate the directions in which research may usefully proceed.

University teachers will probably find this a useful source book but too expensive for individual purchase. The absence of an index is an irritating feature which hinders its use for quick reference by either teachers or students. Essentially it is a collection of informed papers for informed readers. Those who are misled by the title into seeking it out as a 'how to do it' handbook will be doomed to disappointment.

SYLVIA SHIMMIN

Br. J. Psychol. (1971), 62, 2, pp. 289–291

Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review)

AMMON, G. *Gruppendynamik der Aggression*. Berlin: Pinel. 1970. Pp. 175. D.M. 9.80.

BERLIN, B. & KAY, P. *Basic Color Terms: Their Universality and Evolution*. London: University of California Press. 1969. Pp. xi + 178. £4.05.

BEM, D. J. *Beliefs, Attitudes, and Human Affairs*. London: Prentice-Hall. 1970. Pp. x + 114. £1.00.

BERNHARDT, K. S. *Being a Parent: Unchanging Values in a Changing World*. Toronto: Toronto University Press. 1970. Pp. x + 188. £3.60; paper, £1.20.

BERSCHEID, E. & WALSTER, E. H. *Interpersonal Attraction*. London: Addison-Wesley. 1969. Pp. 129. 85p.

BLEULER, E. *Autistic Undisciplined Thinking in Medicine and How to Overcome it*. Darien, Conn.: Hafner. 1970. (Translation of 1963 edition.) Pp. xix + 218. Price not given.

BRADBURN, N. M. *The Structure of Psychological Well-Being*. Chicago: Aldine. 1970. Pp. xvi + 318. $9.75.

BRILL, L. & LEIBERMAN, L. *Authority and Addiction*. London: Churchill. 1969. Pp. xvi + 318. £5.75.

BROWN, F. G. *Principles of Educational and Psychological Testing*. Hinsdale, Ill.: Dryden Press. 1970. Pp. ix + 468. £4.70.

CARSON, R. C. *Interaction Concepts of Personality*. London: Allen & Unwin. 1970. Pp. xiv + 306. £3.00.

CARTER, C. H. *Handbook of Mental Retardation Syndromes*, 2nd edition. Springfield, Ill.: Thomas. 1970. Pp. xvi + 285. $12.75.

CHIN, R. & A. *Psychological Research in Communist China 1949–1966*. London: M.I.T. Press. 1969. Pp. xii + 274. £4.65.

CLARK, K. E. & MILLER, G. A. (eds.). *Psychology*. London: Prentice-Hall. 1970. Pp. ix + 146. £3.00; paper, £1.00.

COOK, P. E. (ed.). *Community Psychology and Community Mental Health: Introductory Readings*. San Francisco: Holden-Day. 1970. Pp. xiii + 286. $10.50.

DAY, R. H. *Human Perception*. Chichester: Wiley. 1970. Pp. xiii + 193. £2.50; paper, £1.25.

DEBUSKEY, M. (ed.). *The Chronically Ill Child and his Family*. Springfield, Ill.: Thomas. 1970. Pp. xvi + 203. $11.50.

DELGADO, J. M. R. *Physical Control of the Mind: Toward a Psychocivilized Society*. New York: Harper & Row. 1969. Pp. xxi + 280. $7.95.

DIENES, Z. P. & JEEVES, M. A. *The Effects of Structural Relations on Transfer*. London: Hutchinson. 1970. Pp. 148. £2.00.

EDELSON, M. *The Practice of Sociotherapy: a Case Study*. London: Yale University Press. 1970. Pp. 345. £5.65.

ERTEL, S. *Psychophonetik: Untersuchungen über Lautsymbolik und Motivation*. Göttingen: Hogrefe. 1969. Pp. 230. D.M. 42.

EVANS, I. M. & SMITH, P. A. *Psychology for a Changing World*. Chichester: Wiley. 1970. Pp. xiii + 444. £3.50.

FESTINGER, L. *A Theory of Cognitive Dissonance*. London: Oxford University Press. 1970. (First published 1957.) Pp. xi + 291. £3.60.

[illegible], C. J. (ed.) *The Future of Psychotherapy*. Boston, Mass.: Little, Brown. 1969. Pp. xii + 415. £3.75.

[illegible], F. *L'orientation professionnelle à l'âge des ordinateurs*. Neuchâtel, Suisse: [illegible] et Niestlé. 1970. Pp. 261. Price not given.

[illegible], J. (ed.) [illegible] *Symposium* Volume 17. *Psychology of Human* [illegible]. London: [illegible]. 1970. Pp. xvi + 235. £2.10.

Greenbaum, S. & Quirk, R. *Elicitation Experiments in English: Linguistic Studies in Use and Attitude*. Harlow, Essex: Longmans. 1970. Pp. 153. £1.75.

Heber, R. *Epidemiology of Mental Retardation*. Springfield, Ill.: Thomas. 1970. Pp. [illegible] + 116. $6.75.

Heim, A. W. *The Appraisal of Intelligence*. London: N.F.E.R. 1970. (First published 1954.) Pp. 169. £1.25.

Hersher, L. (ed.) *Four Psychotherapies*. London: Butterworth. 1970. Pp. xii + 152. £2.00.

Horowitz, I. L. (ed.) *Sociological Self Images: a Collective Portrait*. Oxford: Pergamon Press. 1970. Pp. 255. £3.75.

Houts, M. *They Asked for Death*. New York: Cowles. 1970. Pp. xiii + 241. $6.95.

Jahoda, G. *The Psychology of Superstition*. Harmondsworth, Middx: Penguin Books. 1970. Pp. 158. 30p.

Leplat, J., Enard, C. & Weill-Fassina, A. *La Formation par l'apprentissage*. Paris: Presses Universitaires de France. 1970. Pp. 200. 13 FF.

Levine, S. & Elzey, F. F. *A Programmed Introduction to Educational and Psychological Measurement*. Belmont, Calif.: Brooks/Cole. 1970. Pp. viii + 224. Price not given.

Levis, D. J. (ed.) *Learning Approaches to Therapeutic Behavior Change*. Chicago: Aldine. 1970. Pp. ix + 262. $8.95.

McCammon, R. W. *Human Growth and Development*. Springfield, Ill.: Thomas. 1970. Pp. xi + 295. $9.00.

McGlade, F. S. *Adjustive Behavior and Safe Performance*. Springfield, Ill.: Thomas. 1970. Pp. xi + 198. $9.75.

MacKenzie, N., Eraut, M. & Jones, H. C. *Teaching and Learning: an Introduction to New Methods and Resources in Higher Education*. London: H.M.S.O. (UNESCO, Paris.) Pp. 209. £1.05.

Mannoni, M. *The Child, his 'Illness', and the Others*. London: Tavistock Publications. 1970. (French edition 1967.) Pp. xvi + 286. £3.15.

Maslow, A. H. *Motivation and Personality*, 2nd edition. London: Harper & Row. 1970. (First published 1954.) Pp. xxx + 369. £2.30.

Miller, G. A. *The Psychology of Communication: Seven Essays*. Harmondsworth: Penguin Books. 1970. (First published 1967, U.S.A.) Pp. 192. 30p.

Minium, E. W. *Statistical Reasoning in Psychology and Education*. Chichester: Wiley. 1970. Pp. xx + 465. £4.40.

Pesso, A. *Movement in Psychotherapy: Psychomotor Techniques and Training*. London: University of London Press. 1969. Pp. xii + 221. £3.25.

Piaget, J. & Inhelder, B. *Mental Imagery in the Child: a Study of the Development of Imaginal Representation*. London: Routledge & Kegan Paul. 1971. (1966 in French.) Pp. xix + 396. £3.50.

Przeworski, A. & Teune, H. *The Logic of Comparative Social Inquiry*. Chichester: Wiley. 1970. Pp. xiv + 153. £4.00.

Raknes, O. *Wilhelm Reich and Orgonomy*. Oslo: Universitetsførlaget. 1970. Pp. 130. N.Kr. 24.00.

[illegible] *[illegible] Programming of Children with [illegible]*. Springfield, Ill.: Thomas. 1970. Pp. xii + 123. $7.50.

[illegible] F. A. & [illegible] J. G. *Introduction to Mathematical Psychology*. London: Addison-Wesley. 1970. Pp. xii + 371. £5.15.

[illegible] J. P. *[illegible] Developing Total Talent*. Springfield, Ill.: Thomas. 1970. Pp. [illegible] + 339. $17.50.

[illegible] M. *The [illegible] Impaired School Leaver*. London: University of London Press. 1970. Pp. 240. £1.75.

[illegible] G. F. M. & Walton, H. J. (eds.). *The Training of Psychiatrists*. Conference Proceedings. London: Royal Medico-Psychological Association. 1970. Pp. 15[illegible]. £1.[illegible]

[illegible] W. A. (Jr. ed.). *Personality and Religion: the Role of Religion in Personality Development*. London: S.C.M. Press. 1970. Pp. ix + 245. £1.50.

[illegible] *Collection of Moments: a Study of Involuntary Memories*. London: Longman. 1970. Pp. viii + 142. £2.00.

[illegible] K. L. *Beliefs and Values*. New York: Holt, Rinehart & Winston. 1970. Pp. x + 159. $2.95.

Schultz, D. P. (ed.). *The Science of Psychology: Critical Reflections*. New York: Appleton-Century-Crofts. 1970. Pp. ix + 402. $5.50.

Schwab, J. J. *Handbook of Psychiatric Consultation*. New York: Appleton-Century-Crofts. 1968. Pp. xii + 318. £3.90.

Seidenberg, R. *Marriage in Life and Literature*. New York: Philosophical Library. 1970. Pp. ix + 307. $5.95.

Stacey, M. (ed.). *Hospitals, Children and Their Families*. London: Routledge & Kegan Paul. 1970. Pp. xii + 188. £2.50.

Staff, V. S. *Seven English Poets and Mysticism*. London: Churches' Fellowship for Psychical Spiritual Studies. 1970. Pp. 56. £0.33.

Storr, A. *Human Aggression*. Harmondsworth: Penguin. 1970. (First published 1968.) Pp. 172. £0.25.

Storrow, H. A. *Outline of Clinical Psychiatry*. New York: Appleton-Century-Crofts. 1969. Pp. x + 535. £4.40.

Ward, D. *Music for Slow Learners*. London: College of Special Education. 1970. Pp. 11. £0.20.

Weinstein, G. & Fantini, M. D. (eds.). *Toward Humanistic Education: a Curriculum of Affect*. New York: Praeger. 1970. Pp. xii + 228. Price not given.

Wenrich, W. W. *A Primer of Behavior Modification*. London: Prentice-Hall. 1970. Pp. 94. £1.15.

Wertheimer, M. *A Brief History of Psychology*. New York: Holt, Rinehart & Winston. 1970. Pp. ix + 163. £1.20.

Zimbardo, P. & Ebbesen, E. B. *Influencing Attitudes and Changing Behavior*. London: Addison-Wesley. 1969. Pp. 148. £0.85.

CHANGE-OVER TO SI UNITS

From January 1972 the journals of the British Psychological Society will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Further information about SI units is contained in the revised *Suggestions to Authors* pamphlet issued by the Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

Authors who refer to physical measures in their papers should now normally use SI units; common units of time (e.g. hour, year) will, of course, persist. Conversion tables will be published during the change-over period, but in the *British Journal of Psychology* these will be restricted to units of length, as others rarely occur. When they do, conversion factors will be stated in footnotes.

REVUE DE PSYCHOLOGIE APPLIQUÉE

Publication Trimestrielle

Directeurs: Dr P. PICHOT et P. RENNES

4ème Trimestre 1970 – Volume 20 – No. 4

SOMMAIRE

Rédaction et Administration:
48 Avenue Victor Hugo, PARIS 16e

Abonnements (un an) France: 20 F; Étranger: 24 F
Le numéro France: 6 F; Étranger: 7 F

Réglements au nom des Éditions du Centre de Psychologie Appliquée

C.C.P. PARIS 5.157.61

Numéro spécimen sur demande

International Journal of Psychology
Journal International de Psychologie

Quarterly

Published by the International Union of Psychological Science

The *International Journal of Psychology* is devoted to cross-cultural comparisons of psychological phenomena in general psychology, child psychology and social psychology. It also deals with international confrontation of theories and methods on which psychologists from various countries differ because of their native culture and training. Oriented towards basic research and theory, it publishes scientific papers in English or in French. In each issue, an International Platform gives news and information on internationally significant problems.

Editor (Scientific papers) *G. de Montmollin*
28 rue Serpente, Paris VI°, France

Associate Editor (International Platform) *G. Nielsen*
University of Aarhus, Denmark

VOLUME 5 1970 No. 4

CONTENTS SOMMAIRE

Annual subscription (4 issues) $8.00, F 44 should be sent to

DUNOD, 92 rue Bonaparte, Paris VI°

The British Journal of Psychology

The British Journal of Medical Psychology

The British Journal of Social and Clinical Psychology

The British Journal of Mathematical and Statistical Psychology

The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology, The British Journal of Medical Psychology and *The British Journal of Social and Clinical Psychology* appear in quarterly parts, the subscription price, payable in advance, being £7.50 per volume post free ($25.000 in the U.S.A.) for *The British Journal of Psychology*, and £6.50 each ($22.00) for *The British Journal of Medical Psychology* and *The British Journal of Social and Clinical Psychology*. Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £6 per volume post free ($18 in the U.S.A.). Subscriptions should be sent to the Society's Agents, John Wright & Sons Ltd., the Stonebridge Press, Bath Road, Bristol BS4 5NU; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number. The subscription price in the U.S.A. is $4.20. Subscriptions should be sent to the Subscription Dept Manager, Associated Book Publishers Ltd, Distribution Centre, North Way, Andover, Hants.

The series of *Monograph Supplements* issued in connexion with *The British Journal of Psychology* is now discontinued. Orders may be sent to any bookseller, or Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor W. Sluckin, Department of Psychology, The University, Leicester, LE1 7RH;

for *The British Journal of Medical Psychology* should be sent to

Dr J. J. Sandler, Room 101, Department of Psychiatry, Institute of Psychiatry, De Crespigny Park, London, S.E. 5;

for *The British Journal of Social and Clinical Psychology* should be sent to

Professor Thelma Veness (Social Psychology Editor), University of London Institute of Education, Malet Street, London, W.C. 1, *or* to
Dr J. G. Ingham (Clinical Psychology Editor), M.R.C. External Staff, Llandough Hospital, Penarth, Glamorgan, CF6 1XW;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Dr A. R. Jonckheere, Department of Psychology University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Professor J. D. Nisbet, Department of Education, University of Aberdeen, Old Aberdeen.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £8.50. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1.50 annually). *The British Journal of Medical Psychology* (£1.50 annually). *The British Journal of Social and Clinical Psychology* (£1.50 annually). *The British Journal of Mathematical and Statistical Psychology* (£1.50 annually), and *The British Journal of Educational Psychology* (£1.50 annually).

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18–19 Albermarle Street, London, W1X4DN.

The British Journal of Psychology *Volume 62, Part 2, May 1971*

CONTENTS

Printed in Great Britain at the University Printing House, Cambridge

Volume 62, Part 3 August 1971

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by W. SLUCKIN

with J. ANNETT, W. S. ANTHONY, J. BROWN, R. CONRAD, N. F. DIXON, E. D. FRASER, J. A. GRAY, I. MARTIN, R. L. REID, H. R. SCHAFFER and P. B. WARR

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB

AMERICAN BRANCH:

32 EAST 57TH STREET, NEW YORK, N.Y. 10022

£2.50 net $8.50

Preparation of Papers for The British Journal of Psychology

In submitting papers for publication in *The British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. **56** (1965) onwards.

A brief summary, suitable for use as an abstract, should precede the article.

Manuscripts should be submitted in duplicate.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation: and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.

Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–248.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *Année psychol.* **50**, 83–96.

World list abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

Subscriptions

The British Journal of Psychology is published quarterly.

Annual subscription (including postage) £7.50 net ($25.00 in the U.S.A.).

Second class postage paid at New York, N.Y.

Published by Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB and 32 East 57th Street, New York, N.Y. 10022.

Br. J. Psychol. (1971), 62, 3, pp. 293–302
Printed in Great Britain

MICHOTTE'S EXPERIMENTAL METHODS

By R. B. JOYNSON
Department of Psychology, University of Nottingham

Michotte's account of his experiments on the perception of causality is incomplete and imprecise. There is considerable evidence, however, that subjective variables were very poorly controlled. It follows that little reliance can be placed on his conclusions.

Few experimental studies have been so highly praised as those of Michotte on the perception of causality. The original French edition (Michotte, 1946) was hailed as 'one of the outstanding achievements of contemporary experimental psychology ...an epoch-making flavour...a monument of careful and sustained investigation' (Oldfield, 1949, p. 104). The English translation (Michotte, 1963), which contained additional material, was 'one of the most important and valuable contributions to psychology in the last twenty years' (Vernon, 1964, p. 74). However, in a repetition of some of Michotte's basic experiments, Beasley has obtained very different results, and he concludes that 'there would seem to be a need for a further analysis of Michotte's experimental methods and his theoretical framework' (Beasley, 1968, p. 406). This paper examines Michotte's experimental methods.

Michotte (1963) reports 102 experiments in all. Nos. 1 and 2 are 'type-experiments' concerned with the basic impressions of mechanical causality: 'launching' and 'entraining'. Nos. 3–72 examine the stimulus-conditions in which impressions of launching and entraining occur. Nos. 73–90 concern qualitative causality. All these experiments (nos. 1–90) appear in Michotte (1946), and are reprinted in Michotte (1963). Of the remainder, nos. 91–95 are on 'trace-making'. These appear in Appendix I of Michotte (1963), and originated in 1954. Nos. 96–102 are again on mechanical causality. These appear in Appendix II of Michotte (1963), and originated in 1961.

Michotte's basic contention is that impressions of causality are 'directly experienced' or 'given' (p. 21; page references alone refer to Michotte, 1963); and that they are an immediate response to definable configurations of stimuli, rather than an outcome of past experience or intuition. He states that he uses two sources of information: the experimental situation or system of stimulus-conditions, and the verbal responses of the subjects (p. 305). This paper accepts that the stimulus-conditions were adequately controlled, but provides evidence that factors connected with the subjects and their reports were neglected or mishandled. In evaluating this aspect of Michotte's experiments, particular attention will be paid to the following criteria, which would perhaps be generally accepted. (1) An accurate record of subjects should be made, and the number of subjects should be sufficient to ensure representative results. (2) Subjects should be naive in the sense of being unfamiliar with the literature on the topic, and especially with the experimenter's previous results and present expectations. (3) Subjects should not be examined in large groups under conditions in which suggestion, either from the experimenter or from each other, might bias the results. (4) Where there are numerous experimental conditions, there should be an adequate system of control. (5) Instructions should be clearly stated,

and should be recorded before, not after, results are presented. (6) Subjects' reports should be recorded verbatim, and should be clearly separated both from the experimenter's account of what he thought they meant and from his theoretical interpretations.

MICHOTTE'S INTRODUCTION

In an introductory section, Michotte describes the two type experiments. They are termed the 'fundamental experiments' (p. 19), and are intended to establish the basic contention that the causal impression is 'directly experienced' (p. 21). There is a striking contrast between the wealth of detail concerning the apparatus and the dearth of information about the subjects and the procedure. Michotte says little except that these two experiments 'have been tried out on a large number of subjects (several hundreds) of all ages. All of them have given similar descriptions, with the exception of one or two, who, observing in an extremely analytical way, said that they saw two successive moments, simply co-ordinated in time' (p. 20). There is no further information about who the subjects were, or how exactly they were examined. The statement that the subjects were 'of all ages' might lead some readers to suppose that children were included, especially since research on children would be particularly relevant to Michotte's theories; but one eventually discovers that 'plans for such research have not yet advanced beyond the project stage' (p. 255). It is particularly important, too, that there should have been 'one or two' subjects who observed in 'an extremely analytical way', for these are the subjects who disagreed with his basic contention. But Michotte does not explain how he contrived to detect this negligible minority among the overwhelming majority who agreed with him; nor why, having detected them, it was impossible to count them accurately; nor what criteria he used to determine that their observation was 'analytical'; still less, how he decided that it was 'extremely analytical'. Lastly, in the account of these two fundamental experiments, there is no statement of instructions; there is no indication of how the reports were collected; there is not, so far as I can find, a single verbatim report from any subject; and there is certainly no report which is ascribed to any particular subject.

At the end of this introductory section, however, Michotte gives further information about the subjects used in the experiments in general. He states (p. 40): 'The permanent staff of the laboratory, consisting of Professors Montpellier, Nuttin, and Michotte, as well as the assistants, took part in all of them. Some experiments, which required a large number of subjects, were also tried out on students. Others again – and these were of course the most important ones – have been repeated with very many observers (several hundreds) either during lectures given to groups of university colleagues, or in practical demonstration classes.' There are several disquieting features. Not only is a group (apparently small) of experienced subjects to be used throughout, but this group is to contain Michotte himself, two close colleagues who can hardly have been unaware of his theoretical views, and laboratory assistants who can hardly have been in a position to maintain, over a long series of experiments, a genuine independence from the views of the head of the laboratory. Again, the claim that the most important experiments were conducted on 'several hundreds', while at first sight impressive, loses its force when one realizes that they

were carried out during lectures and demonstrations. Michotte does not explain how it was possible, under these circumstances, to record responses. If a questionnaire was used, it is nowhere mentioned. Nor does he explain how it was possible to avoid suggestion, either from the content of the lecture, or from conversation among the audience, or from such prior knowledge as it would be reasonable to postulate in those who elected to attend these public performances. We are left to speculate whether the 'several hundreds', who were mentioned in the type experiments, also took part during lectures and demonstrations. Presumably they were, since Michotte states that the most important experiments were conducted in this fashion. Certainly this would explain both how it was possible to attain such large numbers, and why the precise figure is not given. However this may be, the combined magnitude and imprecision of Michotte's figures inevitably prompt the questions whether he knew how many subjects there were, who they were, or, indeed, what they thought or said.

Perhaps the most important point which emerges is that Michotte does not seem to have been aware, initially, of the need to guard against the possible effects of prior knowledge and suggestion. It may be that these effects can never be eliminated, but Michotte seems to have been, at the outset, heedless of the dangers. This is shown by his proposal to use a small group of highly sophisticated subjects throughout the experiments; by the fact that other, fresh, subjects seem to have been intended to increase numbers rather than to secure naivety; by his willingness to examine these subjects during lectures and demonstrations; and by the lack of any mention of the problem. Michotte's 'fundamental experiments', then, can hardly be regarded as convincing.

Concerning the subjects

The remaining 100 experiments will now be considered, first with respect to the subjects, and then with respect to the procedure. Beasley (1968, p. 405) writes that Michotte 'is most unspecific in reporting the numbers of subjects on which he carried out his experiments... The amazing fact is that no details whatsoever are given about the subjects used in over 50 per cent of his experiments and in another 25 per cent only experienced subjects were used'. This is no exaggeration; indeed, it will be seen that the figure of 25 per cent is probably a marked underestimate. Such information as the writer has been able to gather, about the numbers of subjects and whether they were experienced or new, is given in Table 1. In view of Michotte's incomplete and unsystematic account, the figures must be regarded with caution.

It will be seen that in 53 experiments (i.e. 53 per cent) no specific information is given either about numbers or degree of experience. In 23 per cent, only experienced subjects are mentioned. In 16 per cent, both experienced and new subjects are included. In 8 per cent, only new subjects are mentioned. When an experiment is classified as 'new only', this does not mean that Michotte specifically states that experienced subjects were not used, but only that he does not specifically state that they were used. But since he has earlier stated that the experienced group, mentioned above, 'took part in all' (p. 40), it may well be that they took part in these experiments too; and that the reason why Michotte does not expressly mention it is simply that, having once said that the experienced group took part in all, he takes it for

granted throughout. If this is so, there would be no experiments involving new subjects only. (It is also true that Michotte does not state that new subjects did not take part in the 'experienced only' experiments; but in this case there are no grounds for supposing that they might have participated, because Michotte nowhere claims that new subjects took part in all.) A further point is that, when Michotte mentions experienced subjects, it seems reasonable to infer that they would be drawn from the experienced group which included Michotte. The main justification for this inference is, of course, Michotte's statement that they took part in all. It is supported by the fact that Michotte nowhere mentions any other source of experienced subjects; that he occasionally gives identifying initials (e.g. 'Mi') which are in accordance with the inference; and that the reported numbers are what might be expected. Finally, the writer would regard as adequate the number of subjects used in the total 24 per cent of experiments in which new subjects are included, but would consider that the mean number (5·6) used in the 23 per cent of 'experienced only' experiments is distinctly unreliable. These include 11 experiments in which the number of subjects is only 3, 2 or even 1.

Table 1. *Estimated distribution of Michotte's experiments among various categories of subject, with estimated mean number of subjects*

Experiments	No specific information	Experienced only	Experienced and new	New only	Total
Launching (3–47)	29	10	2	4	45
Entraining (48–72)	22	1	1	1	25
Qualitative (73–90)	0	11	7	0	18
Trace-making (91–95)	1	0	3	1	5
Mechanical (96–102)	1	1	3	2	7
Total	53	23	16	8	100
Mean number of subjects		5·6	Exp. 4·8 New 18·2	17·6	

What happened in the 53 experiments where no specific information is given? In the absence of further evidence it might be suggested that the subjects would be distributed among the three categories in roughly the same proportion as in the experiments where information is given, namely 23:16:8. But actually there is further evidence. Once again, Michotte has stated that the experienced group took part in all; and once again he may well be assuming that the reader will take it for granted that they participated in these 53 experiments. On the other hand, he has not claimed that new subjects took part in all, and there is no positive evidence to justify the assumption that any were included here. Thus the most likely conclusion is that the subjects in these 53 experiments were drawn from the experienced group, and from this group alone. This is strengthened by the reflexion that it would have been much easier for Michotte, in such numerous experiments, to call upon the experienced group who were ready to hand, than to collect new subjects on each occasion. Thus *the probable distribution for the whole 100 experiments is 76 per cent 'experienced only'* (i.e. 53 'no specific information' plus 23 'experienced only'), *16 per cent 'experienced and new', and 8 per cent 'new only'*. Then experienced subjects would have been included in 92 per cent, perhaps all; and new subjects would have been included in 24 per cent.

Any suggestion that new subjects might have been used more frequently needs to be tempered by a realization of the large numbers involved. The 24 per cent of experiments in which new subjects were definitely used must have taken between 400 and 500. If new subjects were used at the same rate in only half of the 'no specific information' experiments, a further 400 to 500 would have been required. There were in addition 'several hundreds' in the type-experiments. A total of, say, 1500 new subjects is not of course impossible; but it must be remembered that the greater the number claimed, the more onerous becomes the task of examining them all with care, and the greater the assurance needed that care was in fact exercised.

Lastly, it is again reasonable to suppose that the experienced subjects would be drawn from the small expert group, and that their average number would be no greater than the average number used in the 'experienced only' experiments, namely 5·6. Indeed, it may well have been less than this, for in the 77 experiments on mechanical causality (nos. 3–72 and 96–102), which are the most important for Michotte's general conclusions, the number of experienced subjects, where given, averages only 2·75. In short, so far as one can judge from Michotte's economical narrative, *it is likely that a small expert group, which averaged less than six in number and included Michotte as a leading member, formed the sole subjects in some 75 per cent of experiments, and were included in some 90 per cent, perhaps in all.*

Vernon (1964, p. 74) writes: 'It should be noted that Michotte carefully refrained from suggesting to his subjects that they might perceive these causal effects; and in all his later work he used naive subjects who had no expectations as to what they would see.' With respect to the first half of this sentence, it should be noted that for Michotte to refrain from suggesting to himself and a few close associates that they might perceive these causal effects, during a long series of experiments in which they were continually reporting them, is a monument of careful and sustained restraint. With respect to the second half, the phrase 'all his later work' presumably refers to experiments 91–95 (of 1954) and 96–102 (of 1961). In this group of 12 experiments there is in fact no information about who took part in two experiments (nos. 93 and 96); in one experiment (no. 97) only experienced subjects are reported to have been used; and although there are nine experiments in which new subjects were included, there are only three (nos. 94, 100, 101) in which new subjects alone are mentioned. Thus 'in all his later work he used naive subjects' is a highly misleading statement which at best refers to a comparatively small proportion of the total investigation.

At this point it might be urged that Michotte did, after all, use new subjects in at least 24 per cent of his experiments as a whole; that they comprise a very respectable total of between 400 and 500, perhaps more; and that they surely provide a very solid body of evidence to set against the possible effects of suggestion where experienced subjects were used. However, an opinion on the value of the experiments, whether with new or experienced subjects, requires also a study of the experimental procedure.

Concerning the procedure

Under what circumstances were the experiments conducted? No doubt the experienced subjects were examined individually in the laboratory; though this would not eliminate their prior knowledge of the topic, nor the experience and

expectations which they would accumulate during the experiments, nor the discussions which they would in all probability hold among themselves between experiments. In the case of new subjects, the position is obscure. Michotte has said that the most important experiments were conducted during lectures and demonstrations, and in view of the large numbers, this would certainly have been easier than examining them all individually. The only thing of which we can be sure is that we do not know under what circumstances the new subjects were examined; but we cannot exclude the possibility that a substantial proportion took part on these public occasions, with all the difficulties of avoiding suggestion and collecting reports which have already been noted.

What were the instructions? The first statement of instructions which I have been able to find appears in Appendix II, which was written in 1961, some 15 years after 90 per cent of the experiments had been published. Michotte here writes as follows: 'Their instructions were, "Say simply what is going on in the apparatus" or some similar wording such as "Say what you see in the apparatus"; and to make the position clearly understood they were told also that the description should be such as one might give when witnessing any sort of event in ordinary life, and should be given in as spontaneous a way as possible' (p. 305). Why does Michotte mention a variety of possible wordings? Is it because no regular instructions were used? Or because different instructions were used in different experiments? Or because Michotte has mislaid his records and cannot quite remember what they were? Where naive subjects were used, we shall probably never know. But where the subjects were experienced, it seems likely that no consistent wording was used, for instructions can hardly have figured very prominently in the proceedings, at any rate after the first occasion, and least of all when Michotte himself was a subject. It may be added that precise and consistent instructions are of the first importance, especially in view of the multiple meanings of words such as *see*. It is by no means clear that to say what is seen in the apparatus, and to say what is going on in it, are identical tasks; yet possible ambiguities might be decisive in experiments where conclusions are to be drawn about what is given in direct experience.

Michotte (p. 305) continues: 'When we wished to obtain fuller information, the only questions which we allowed ourselves to ask were ones such as "Could you not put it another way?", "Could you not be a little more precise?" or "Could you not give me a few more details?" with all traces of suggestion carefully avoided.' This passage certainly demonstrates that Michotte was alive to the need for avoiding suggestion in 1961, during his last seven experiments. But the crucial question is whether he had been sufficiently aware of the problem so many years before, when the majority of his experiments were conducted. Michotte seems convinced that he had been, but his conviction is so obviously conducive to his own interests, and recorded so long after the event, that one must ask for independent evidence. But we have already found evidence, mentioned above, that he did not seem to be aware of the problem initially; and our confidence has been further undermined by discovering how large a part was in fact played by the small expert group. Further, it can hardly be denied that there is a certain implausibility in the suggestion that the above questions were used very extensively with this group, though in the case of Michotte himself the request to be a little more precise, or give a few more details, would not

have lacked its point. Where naive subjects were used, it would have been very difficult to administer these questions during lectures; though if they were examined individually, why are no examples given of the kinds of occasion on which fuller information was desired, or of the kinds of answer obtained? Why is there also no indication of how many subjects were examined in this way – whether 'one or two' or 'several hundreds'? There seems to be very little evidence that this procedure was followed at all widely. Yet the onus of proof that experiments have been properly conducted rests squarely upon the experimenter. He cannot hope to carry conviction with belated assurances of his proficiency; indeed, the more loudly he bangs the stable door, the more inclined we are to ask what has happened to the horse.

The most important questions concern the subjects' reports. The experimental details which have been considered so far are only means towards the end of securing reliable and representative reports. Even if all those details were in order, the experiments would still be of little value unless it could also be demonstrated that the collection and treatment of the subjects' reports had been satisfactorily performed. Concerning this, the following points need to be made. (1) In view of the large number of experiments, and especially of the large number of naive subjects, it would be natural to anticipate a wealth of examples culled from the voluminous records which must presumably have accumulated. It is therefore both disappointing and puzzling to discover that verbatim records of subjects' reports are remarkably sparse, and particularly so for naive subjects. We have already noted that there are no verbatim reports for nos. 1 and 2. In nos. 3–72, they are thinly scattered. There are rather more in nos. 73–95. In the final experiments on mechanical causality, nos. 96–102, I can find only one verbatim report, and that concerns the description of a bicoloured rectangle, 'or, as one observer imaginatively put it, a "small flag"' (p. 328). (2) For the most part, the subjects' reports are given by Michotte *indirectly, and for the experimental group as a whole*. Thus for no. 28: 'The result of this experiment is absolutely clear. The Launching Effect can be quite as evident, quite as "good", as in the most successful experiments described so far. This fact has been confirmed by a large number of people, and in particular by at least a dozen subjects who were experienced scientific observers in different fields' (p. 85). It should be obvious that such indirect summaries are highly unreliable, since they permit the experimenter's own views to colour the conclusions to an entirely unknown extent. In this case, numerous questions arise. How could these people confirm what Michotte calls 'this fact'? Had they also taken part in 'the most successful experiments so far'? How long ago? What did they actually say? Is a successful experiment one in which Michotte's expectations are confirmed? Is the use of 'experienced scientific observers' supposed to impress us? Yet this is a 'result' which Michotte calls 'absolutely clear'. (3) It must always be remembered that Michotte was himself a subject. Thus when he says that 'an observer stated' or 'observers reported', it can easily be assumed that he is necessarily recounting what other people – independent observers – have reported, whereas it is always possible that the observation is his own. (4) Comparison between the reports of experienced and naive subjects is especially important as a control for the effects of experience. Such comparisons are infrequent in Michotte's account, but one's confidence is not increased by his handling of an instance where the reports of four naive subjects conflicted with those of an undisclosed number of

experts. Here Michotte writes that 'we thought it wise, as a matter of principle, to abide by the observations of the experienced subjects' (p. 76). (5) Michotte's informal group reporting is frequently interspersed with his theoretical interpretations, and nowhere are experimental observations and theoretical deductions placed in clearly distinguishable sections. Thus for no. 57: 'Here, too, it is possible to obtain a clear impression of traction, especially at the moment when the second object starts. It is as though the first object were towing the other at the end of a piece of string. The impression is not as good as in the previous experiment. This is undoubtedly because of the separation of the objects in space; the observer has to be prepared to look at the experiment as a whole if the Traction Effect is to occur. An analytical approach, which can creep in very easily, is fatal to it' (p. 161). This extract is, I believe, a fair example of Michotte's method. With no indication of the identity or number of the subjects, he asserts a general result, including a comparison with a previous experiment. In the same breath, a supposedly indubitable explanation is claimed, and further alleged observations at once make their appearance to explain away any possible disconfirmation. There may be many ways of conducting a scientific inquiry, but this is hardly one of them.

Vernon (1964, pp. 74–5) states: 'Particularly significant is his use of the spontaneous verbal reports of naive subjects as constituting evidence on the phenomenal nature of their perceptions.' This would indeed have been very desirable, but there seems little evidence that the compliment is deserved.

DISCUSSION AND CONCLUSIONS

Where experimental methods are unsound, it does not follow that conclusions are false, only that they are uncertain. A fair comment on Michotte's experiments seems, therefore, to be one which he makes himself upon the work of another experimenter (Tognazzo) whose results differed from his own. Michotte writes that 'in the absence of more complete data, it is impossible to reach a valid conclusion as to the significance of these experiments. To be convincing they should, in my opinion, have been carried out on subjects who were new to the task but familiar with scientific observation; also they should have been performed in isolation, i.e. not forming part of a series in which other causal situations were presented, and in addition one would need to be quite sure what exactly a subject meant by some particular response' (p. 337, note).

It would not be surprising if other investigators failed, in whole or in part, to confirm Michotte's claims. Gruber *et al.* (1957) found evidence for the effect of experience. Gemelli & Capellini (1958) found much greater proneness to make analytic judgements. Powesland (1959) found practice important. Boyle (1960), who worked under the direction of Michotte, found it necessary to administer a preliminary screening test in which 50 per cent of his subjects were rejected 'either because their replies were completely ambiguous, or because they described the stimuli in ways opposite to that expected' (Boyle, 1960, p. 173). This technique deserves to be more widely known among experimenters who find that their subjects fail to give the right answers. Beasley (1968), in a close repetition of the launching and entraining experiments, found that only about one third of his naive subjects gave causal

responses (12 of 34 and 13 of 34 respectively). Also in experiments using film loops some 88 per cent of 181 naive subjects alternated between causal and non-causal responses, and the percentage of causal responses in the launching and entraining situations was 61 and 45 respectively. It seems likely therefore that as many as half of a group of naive subjects would fail to report a causal impression in Michotte's basic situation, instead of only 'one or two' out of 'several hundreds' as he supposed.

But it may be felt that the positive value of Michotte's contribution has been overlooked. It might be said that Michotte has, most ingeniously, demonstrated the possibility of an experimental approach in a new and difficult field; that no account has been taken of the significant ideas which he has introduced; and that the upshot is simply that Michotte has, in some respects, been overtaken by the natural, and by no means inglorious, fate of the pioneer—those who come after him can see how to improve upon his work.

There would be much justice in such a rejoinder, though Michotte's theoretical contribution cannot be considered here. But when he is credited with introducing experimental method in the field of causality, this needs to be qualified by reflexion upon the general limitations of experiment in psychology. Bartlett long ago observed that there was a tendency for experimental psychologists to place the burden of explanation upon the stimulus; and he added that 'to make the explanation depend mainly upon variations of stimuli...is to ignore dangerously those equally important conditions of response which belong to the subjective attitude and to predetermined reaction tendencies' (Bartlett, 1932, p. 4). This tendency, which is as prominent in the Behaviourist as in the European tradition, is clearly exemplified by Michotte. It is shown in his heavy emphasis on his apparatus for the control of external stimuli; in his belief that the causal impression is an immediate response to those stimuli; and in his relative neglect of the subjects and their reports. But if Bartlett was right, experimental work which overlooks inner conditions is of severely limited value.

Unfortunately it remains true that it is extremely difficult to elucidate and control inner conditions. The attempt to do so raises acute methodological problems, and the writer has elsewhere urged the weakness of some widely advocated solutions (Joynson, 1970). The implication of the present paper is that much greater attention should be given to subjects' reports. This certainly has its own difficulties. The variety of individual reactions, even in relatively simple perceptual situations, is very great; and it is hard to interpret correctly the subtle shades of meaning which abound (e.g. Joynson & Newson, 1962). Such data are notoriously resistant to conventional scientific treatment, though this does not mean that they are either non-existent or unimportant.

A purely 'objective' approach is sometimes thought to be the answer. It would of course be possible to ask the subject not to describe his perceptual impressions but to choose between alternative motor responses according to whether he had, or had not, perceived causality. But it is obvious that this only shifts the problem superficially. The procedure would still require verbal instructions, and would seem to assume that phrases such as 'perceive causality' have a plain and agreed meaning. But it is precisely such meanings which need to be elucidated. The procedure skirts this issue by assuming that experimenter and subjects know perfectly well what is

meant by 'perceiving causality'. If they do know this, it is not easy to see why they should not keep to ordinary language. If they do not know it, the procedure is ridiculous. At this point, however, a discussion of experimental method becomes an argument about theoretical framework.

REFERENCES

BARTLETT, F. C. (1932). *Remembering*. Cambridge University Press.

BEASLEY, N. A. (1968). The extent of individual differences in the perception of causality. *Can. J. Psychol.* **22**, 399–407.

BOYLE, D. G. (1960). A contribution to the study of phenomenal causation. *Q. J. exp. Psychol.* **12**, 171–179.

GEMELLI, A. & CAPPELLINI, A. (1958). The influence of the subject's attitude in perception. *Acta psychol.* **14**, 12–23.

GRUBER, H. E., FINK, C. D. & DAMM, V. (1957). Effects of experience on perception of causality. *J. exp. Psychol.* **53**, 89–93.

JOYNSON, R. B. (1970). The breakdown of modern psychology. *Bull. Br. psychol. Soc.* **23**, 261–269.

JOYNSON, R. B. & NEWSON, L. J. (1962). The perception of shape as a function of inclination. *Br. J. Psychol.* **53**, 1–15.

MICHOTTE, A. (1946). *La Perception de la causalité*. Louvain: Institut Supérieur de Philosophie.

MICHOTTE, A. (1963). *The Perception of Causality*. London: Methuen.

OLDFIELD, R. C. (1949). Review of Michotte (1946). *Q. J. exp. Psychol.* **1**, 104–105.

POWESLAND, P. F. (1959). The effect of practice upon the perception of causality. *Can. J. Psychol.* **13**, 155–168.

VERNON, M. D. (1964). Review of Michotte (1963). *Br. J. soc. clin. Psychol.* **3**, 74–75.

(*Manuscript received* 19 *May* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 303–309
Printed in Great Britain

VERBAL TRANSFORMATIONS ON REPEATED LISTENING TO SOME ENGLISH CONSONANTS

By J. M. CLEGG

Division of Computer Science, National Physical Laboratory
Teddington, Middlesex

Verbal transformations of common English consonants were analysed according to five linguistic features. A consonant and its transform tended to share the features of voicing, nasality and affrication, but not of duration and place of articulation.

The verbal transformation effect (VT) is the name given by Warren (1961 *a*) to the subjective changes experienced when listening to a word presented repeatedly using a loop of recorded tape. The phenomenon was discovered by Warren & Gregory in 1958, and subsequent work on it has been reviewed by Warren (1968). Factors investigated include monotic *v.* diotic stimulation (Warren 1961 *a*), complexity of stimulus words (Warren, 1961 *b*), age differences in the numbers of VTs experienced (Warren, 1961 *b*, 1962; Warren & Warren, 1966), bias resulting from instructions to subjects (Taylor & Hemming, 1963), meaningful *v.* nonsense stimuli (Natsoulas, 1965), familiarity of stimuli (Evans & Kitson, 1967) and context *v.* no context for stimuli consisting of word pairs (Clegg & Grover, 1968).

With this concentration on high-level personal and linguistic factors and the way they affect VTs, little attention has been paid to the nature of the phenomenon itself. Barnett (1964; cited by Warren, 1968), however, analysed phonetically the responses to 12 words. She noted a stability for the voiced/voiceless property of consonants and for the type of change characteristic of the consonants and vowels occurring in her words.

The present study was done to find out how consonants transform, and to compare transformations with perceptual confusions, which were investigated by Miller & Nicely (1955). They published a comprehensive set of confusion matrices showing the frequency of each kind of error in perceiving 16 English consonants under 17 conditions of signal/noise ratio and high and low frequency cut-off. Miller & Nicely analysed their results in terms of the following five linguistic features.

(1) *Voicing*. The vocal chords vibrate during articulation of the consonants /b d g v ð z ʒ m n/, but not of /p t k f θ s ʃ/.

(2) *Nasality*. In articulating /m/ and /n/ the nasal cavity provides a resonance which is not used for the other consonants.

(3) *Affrication*. If the articulators close completely, the consonant may be a stop or nasal, but if they are brought close together and air is forced between them, the result is a kind of friction noise which comes with /f θ s ʃ v ð z ʒ/ but not with /p t k b d g m n/.

(4) *Duration*. An arbitrary distinction was made between /s ʃ z ʒ/ and the rest because of the longer time taken to articulate them.

(5) *Place of articulation*. This refers to the part of the mouth where the major constriction of the vocal tract occurs: front (/p f b v m/), middle (/t d θ s ð z n/) and

back /k g ʃ ʒ/. Miller & Nicely considered this the least satisfactory and most superficial of their five features.

Articulation scores for each feature were calculated by, for example, grouping the voiced consonants v. the unvoiced, and from the data estimates were made of the probability that the voicing feature would be perceived correctly. In all listening conditions all features were guessed or heard correctly with greater frequency than would have been expected from random guessing, although at some extreme conditions this difference was fairly small.

The consonants used in the present study were /p t k f θ s ʃ b d g v ð z m n l r h/. This set differs from Miller & Nicely's in the omission of /ʒ/, since it does not occur as an initial consonant in English, and the addition of /l r h/, which were included to provide general information on VTs. In not also adding /w y č ǰ/ I followed Miller & Nicely's perhaps now outdated view that /w y/ are vowels and that /č ǰ/ are the consonant clusters /dʒ tʃ/.

METHOD

In order to enunciate most consonants it is necessary for there to be a vowel immediately next to them. To confine VTs to the consonants, an unproductive vowel was chosen to follow each consonant—the vowel /i/, as in 'feet', since B. E. Pay (1968, personal communication) found it to be the least productive of transformations compared with nine of the other most frequently used English vowels in a /b–vowel–t/ context ('bat', 'boot', etc.).

The speakers for the stimulus tapes were five men aged 23–37, four of whom had spent their lives in the London area, the other having a 'B.B.C.' accent. They spoke each of the 18 consonants followed by /i/ into a single-track Nagra tape-recorder (15 in./sec.) at similar loudnesses in a soundproof chamber. Each speaker spoke his list in a different random order to avoid systematic sequential changes in tone of voice. The recordings were made into loops 11·42 in. long, which was about 0·5 in. longer on the tape than the longest of all the 90 recorded sounds. The shortest sound had 7·1 in. of blank tape. These loops were then played on the Nagra at a fixed volume and re-recorded on to conventional reels for 400 cycles (about 4½ min.) on a Ferrograph twin-track tape-recorder run at 7½ in./sec. The input gain on the Ferrograph was adjusted to make the recorded level constant for each loop. Afterwards the S/N ratio was measured for a random sample of 31 of the 90 recordings, giving a mean of 30·6 db with a standard deviation of 3·2 db.

In several previous studies of VTs, responses have been obtained by having the subjects speak out their VTs as they 'hear' them. This method produces auditory stimulation in addition to the repeating sounds. Accordingly, VTs experienced in this study were indicated on sheets containing the sounds they were likely to hear (i.e. all the 18 stimuli) written as pseudo-English words, e.g. 'fee', 'kee'. It is often felt when observing the VT effect that one can sometimes consciously reproduce VTs, apparently provided that they have already been experienced. In case previous experience is not a condition for conscious influence, and in case looking at a 'word' might tend to influence that word's occurrence as a transform, any such possible effect was randomized by having the response sheets list the sounds in four different random orders; the various sheets were used in rotation. There was extra space for transformations not listed.

The subjects were five males aged between 18 and 26. They were given the following written instructions: 'You are to sit wearing the earphones. You will hear a voice repeatedly saying one of the English consonant sounds followed by the vowel sound "ee". Now most people in these conditions find that after a time the voice appears to be saying something different from what it was saying at the start. We are interested in what these changes are for a variety of consonantal sounds. You will have a sheet on which to write ticks next to those sounds which the voice seems to be saying. Any changes in the vowel part are of no interest to us—we want you to indicate just the consonant changes. If you hear sounds which are not close enough to what is written on your sheet, please write them at the bottom, under "others". You need write one tick for each sound you hear, no matter how often you hear it. Please take care with the two sounds represented in writing by "thee" such as occur in "thin" and "this", and remember that "gee" does

[illegible] can "per". You will hear two beeps followed by 5 sec. of silence before the voice starts. Each [illegible] lasts 4½ min.'

The earphones were Sharpe HA 10 circumaural, and the Ferrograph was used at a constant level at the ear of 67 db (re 2×10^{-5} N m^{-2}).

Each subject heard the whole series of consonants twice, each series being in random order, with different random orders for each subject. Thus each subject heard 36 stimuli. Each consonant [illegible] each speaker was used twice; each listener heard each voice as nearly as possible with equal frequency, no subject heard the same voice twice with the same consonant tape. Subjects came twice per day to take part, and heard one tape on each occasion. They were not told in advance which sound they would be hearing.

Results and discussion

The data for all subjects have been pooled for the 10 sessions with each consonant and are shown as a transform matrix in Table 1. The stimulus consonants are listed in the left-most column, and the verbal transforms are indicated across the top of the table. The number in each cell is the number of sessions (maximum possible = 10) during which each stimulus-transform pair occurred.

Table 1. *Transform matrix*

Stimuli	p	t	k	f	θ	s	ʃ	b	d	g	v	ð	z	m	n	l	r	h	Vanish
	Consonantal transformations																		
p	–	*3*	*2*	.	*1*	.	.	.	.	.	.	.	.	.	.			5	
t	*5*	–	*1*	.	.	.	.	*1*	.	.	.	.	.	.	.			7	
k	*5*	*2*	–	.	.	.	.	*3*	*1*	*3*	.	.	.	.	.			7	1
f	*5*	.	.	–	*5*	*1*	*1*	*6*	.	.	*3*	*2*	.	.	.			1	1
θ	*7*	*1*	.	*10*	–	*4*	.	*6*	*3*	.	*2*	*3*	*1*	.	.			3	1
s	*6*	*5*	.	*3*	*1*	–	*1*	*5*	*1*	.	*3*	*1*	*3*	.	.			1	1
ʃ	*4*	*2*	*3*	*1*	.	*3*	–	*3*	*2*	*3*	*1*	.	*1*	.	.			3	1
b	*3*	.	.	.	.	.	.	–	*3*	.	*1*	.	.	.	.				2
d	.	.	.	.	.	.	.	*7*	–	*1*	*1*	.	.	.	.			1	3
g	.	*1*	*4*	.	.	.	.	*1*	*3*	–	.	.	.	.	.			1	2
v	*1*	.	.	*3*	*1*	.	.	*3*	.	.	–	*6*	*1*	.	.	1			1
ð	*1*	.	.	*2*	*2*	.	.	*3*	*1*	.	*9*	–	*2*	.	.	2	1		
z	*1*	.	.	.	.	.	.	*1*	*4*	.	*6*	*6*	–	.	.	1		1	
m	.	.	.	.	.	.	.	*6*	.	.	*1*	*1*	.	–	*5*	2			
n	.	.	.	.	.	.	.	*3*	*1*	.	*1*	*1*	.	*6*	–	2		1	3
l					1			4	3		2	2			1	–			
r																	–		1
h	6		2					3										–	

The column heading 'vanish' refers to the total disappearance of the consonant. The consonants /l r h/, which were not used by Miller & Nicely, are omitted from the comparison with their data. In addition, it may be seen from Table 1 that /r/ gave only one transform, and was itself a transform only once. /h/ could be considered acoustically nearly identical to 'vanish', and by adding /l/ to the analysis considering it as voiced, non-nasal, non-affricative, short and articulated in the middle, the conclusions that follow were not invalidated. The data considered in the remainder of this paper are printed in italic type in Table 1.

Following the method of Miller & Nicely, the consonants were grouped according to the presence or absence of the five linguistic features.

This is given in Table 2, together with the statistical analysis following the method which Klatt (1968) used on Wickelgren's (1966) data on confusions of consonants in

short-term memory. The expected values for each cell in the matrix of Table 1 were calculated, accounting for the different productivities of stimuli and the different 'popularities' of VTs. The total number of expected frequencies was equated to the total observed. Using Klatt's notation, a submatrix of the VT matrix is taken, in which the frequencies are grouped according to the presence of a feature (+F) or its absence (−F) in stimulus or VT. See Fig. 1. N_{cs} is the number of cross-set transforms, where $N_{cs} = N_{-+} + N_{+-}$. A feature F is indicated by the data if N_{cs} is smaller than expected.

Table 2. *Analysis of cross-set transformations by linguistic feature*

Stimulus feature		Frequency of VT: Observed			Expected			$\hat{N}_{cs}-N_{cs}$	σ	Dev(N_{cs})	P
Voicing		+V		−V	+V		−V				
/b d g v ð z m n/	+V	84		19	56·20		45·49				
/p t k f θ s ʃ/	−V	57		82	85·70		54·61	55·20	7·75	7·12	< 0·001
Nasality		+N		−N	+N		−N				
/m n/	+N	11		14	0·60		25·25				
(the rest)	−N	0		217	10·44		205·71	21·69	5·52	3·93	< 0·001
Affrication		+A		−A	+A		−A				
/f θ s ʃ v ð z/	+A	88		77	59·58		106·06				
/p t k b d g m n/	−A	7		70	31·99		44·36	54·05	7·70	7·02	< 0·001
Duration (L = long)		+L		−L	+L		−L				
/s ʃ z/	+L	8		62	3·66		68·57				
(the rest)	−L	10		162	13·54		156·23	10·10	7·37	1·37	n.s.
Place		F	M	B	F	M	B				
/p f b v m/	Front	32	29	3	32·35	23·51	5·32				
/t d θ s ð z n/	Middle	89	40	3	80·23	42·04	10·97				
/k g ʃ/	Back	18	15	13	27·96	16·90	2·73	7·89			

		VT: +F	−F
Stimulus	+F	N_{++}	N_{+-}
	−F	N_{-+}	N_{--}

Fig. 1. General transform matrix for a feature F.

The standard deviation of the difference between the observed cross-set transform frequency N_{cs} and its expected value $\hat{N}_{cs}$ is given by

$$\sigma = nq(1-q)^{\frac{1}{2}},$$

where n is the total number of VTs, and $q = N_{cs}/n$. Dev(N_{cs}) summarizes the evaluation of a feature:

$$\mathrm{Dev}(N_{cs}) = (\hat{N}_{cs}-N_{cs})/\sigma.$$

For unstructured data the distribution of Dev(N_{cs}) is approximately normal with a standard deviation of σ.

Table 2 contains a summary of this process for all five features except place of articulation, where Klatt's test is inappropriate for a three-valued feature. However, the expected values were calculated in the same way, and are shown in the table. By inspection it is clear that there is no tendency for VTs and stimuli to share the front or middle-feature values. If the place feature is binarized into 'back' and 'not back',

it would appear that place could become a discriminating feature. This was tested and supported: see Table 3. Although the hypothesis tested here is *post hoc*, being generated by the data, another reason independent of the data can support it: the distance inside the mouth between 'back' and 'middle' is greater than the distance between 'middle' and 'front'; therefore if one is to have a binary place feature, 'back' and 'not back' are the two values to use.

It can be seen that there is a statistically significant tendency for a stimulus and its transforms to share the presence or absence of the three features of voicing, affrication, nasality, and place of articulation if binarized into 'back' and 'not back'. Duration appears to be unimportant.

Table 3. *Analysis of cross-set transformations for binary place of articulation*

Stimulus feature and place	Frequency of VT: Observed		Frequency of VT: Expected					
	Back	Not back	Back	Not back	$\hat{N}_{cs} - N_{cs}$	σ	Dev(N_{cs})	P
Back /k g ʃ/	13	33	2·73	44·86				
Not back (rest)	6	190	16·28	178·13	22·14	6·76	3·28	< 0·01

Table 4. *Average numbers of VTs (per possible stimulus–VT pair) sharing features with stimuli*

Subject	No. of features common to stimulus and VT: 0	1	2	3	4
A	0	0	0	0·04	0·27
B	0	0·15	0·20	0·24	0·52
C	0	0·25	0·15	0·25	0·56
D	0	0·10	0·05	0·21	0·73
E	0	0·10	0·08	0·16	0·67
Average	0	0·12	0·10	0·18	0·69
PP	6	20	60	76	48
Test	$P = 43$;	$z = 3{\cdot}92$;	$P < 0{\cdot}01$		

For the second part of the analysis the transformed consonants were arranged in columns for each subject according to how many of the given linguistic features they shared with the stimulus consonant.

Each figure was divided by the appropriate number of possible consonant pairs (PP) which could theoretically form stimulus and VT; the results of this appear in Table 4. The figures were ranked for each subject and tested against the prediction that VTs and their stimuli are likely to share many features rather than few, using the Kendall S test for the extent to which the predicted and observed orders are associated. The prediction was supported ($P < 0{\cdot}01$).

As we have seen, then, there is a definite tendency for the VTs of the consonants used in this study to share the presence or absence of the linguistic features of voicing, nasality and affrication, and tentatively of place of articulation if this is given the binary values of 'back' and 'not back' instead of Miller & Nicely's three values. Likewise, there is a clear tendency for the majority of VTs to share most features with their stimuli.

The similarities with Miller & Nicely's results should be treated with some caution: (1) the figures in Table 1 are the numbers of subjects × sessions in which a particular VT was heard at least once. The exact frequency of occurrence of each VT could be more useful, but it was considered too difficult to get this information unless the subjects spoke out loud during trials; (2) it is not known if or how the different vowels used here (i) and in the confusion study (/a/) affect either VTs or confusions; (3) Miller & Nicely's speakers were American females; here they were British males.

The results described suggest that the phenomena of verbal transformations and consonantal confusions are mediated by the same perceptual mechanism, a mechanism which hears sometimes correctly, sometimes wrongly. Under normal listening conditions if a string of unexpected words have to be identified in the absence of context, some will be misheard; if the same word were to occur many times widely separated by many other words in an extremely long string, the expectation would be that the correct and incorrect perceptions of that word would be the same as the verbal transformations of the same word heard from a loop of tape. Unfortunately, it would be very tedious to test this suggestion because of the gigantic size of string necessary to prevent a subject from noticing that the word being investigated is recurring. The suggestion fails to account for the time spent listening to a VT stimulus tape before the onset of VTs—a latency which presumably has no counterpart in a confusion experiment.

This work was carried out as part of the research programme of the National Physical Laboratory. The author thanks B. E. Pay, C. M. B. Anderson and C. A. Godbold for their help.

References

Barnett, M. R. (1964). Perceived phonetic changes in the verbal transformation effect. (Doctoral dissertation, Ohio University. Ann Arbor, Mich.: University Microfilms, no. 64-7328.)

Clegg, J. M. & Grover, R. P. (1968). Context and the verbal transformation effect. (National Physical Laboratory, Teddington. Report no. Com. Sci. T.M.19.)

Evans, C. R. & Kitson, A. (1967). An experimental investigation of the relation between the 'familiarity' of a word and the number of changes in its perception which occur with repeated presentation as a 'stabilized auditory image'. (National Physical Laboratory, Teddington. Report no. Auto 36.)

Klatt, D. H. (1968). Structure of confusions in short-term memory between English consonants. *J. acoust. Soc. Am.* **44**, 401–407.

Miller, G. A. & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. acoust. Soc. Am.* **27**, 338–352.

Natsoulas, T. (1965). A study of the verbal-transformation effect. *Am. J. Psychol.* **78**, 257–263.

Taylor, M. M. & Hemming, G. B. (1963). Verbal transformations and an effect of instructional bias on perception. *Can. J. Psychol.* **17**, 210–223.

Warren, R. M. (1961*a*). Illusory changes of distinct speech upon repetition—the verbal transformation effect. *Br. J. Psychol.* **52**, 249–258.

Warren, R. M. (1961*b*). Illusory changes in repeated words: differences between young adults and the aged. *Am. J. Psychol.* **74**, 506–516.

Warren, R. M. (1962). An example of more accurate auditory perception in the aged. In C. Tibbits & W. Donahue (eds.), *Social and Psychological Effects of Aging*. New York: Columbia University Press.

Warren, R. M. (1968). Verbal transformation effect and auditory perceptual mechanisms. *Psychol. Bull.* **70**, 261–270.

Warren, R. M. & Gregory, R. L. (1958). An auditory analogue of the visual reversible figure. *Am. J. Psychol.* **71**, 612–613.
Warren, R. M. & Warren, R. P. (1966). A comparison of speech perception in childhood, maturity and old age by means of the verbal transformation effect. *J. verb. Learn. verb. Behav.* **5**, 142–146.
Wickelgren, W. A. (1966). Distinctive features and errors in short-term memory for English consonants. *J. acoust. Soc. Am.* **39**, 388–398.

(*Manuscript received* 25 *April* 1970; *revised manuscript received* 18 *June* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 311–326
Printed in Great Britain

SEMANTICS IN THE PERCEPTION OF VERTICALITY

By WILLIAM G. CHASE AND HERBERT H. CLARK

Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pa., U.S.A., and Department of Psychology, Stanford University, Stanford, Calif., U.S.A.

Seymour has shown recently that people take less time to judge that the word *above* correctly describes the spatial position of a small circle drawn ABOVE a large reference square than they do for the word *below* and the circle BELOW the square. Seymour has attributed this asymmetry to the tendency for people to invariably scan a picture from top to bottom. In the present study, the first experiment confirms Seymour's results, but the next three demonstrate that the asymmetry Seymour found cannot be accounted for by an attentional-scanning process. Instead, it is proposed that people interpret the words *above* and *below* as abstract symbols at a first stage of processing, interpret the pictures ABOVE and BELOW as abstract symbols at an independent second stage, and compare these two sets of symbols at an independent third stage. In support of this model, the results show, for example, that *above* is interpreted about 80 msec. faster than *below* at the first stage quite independently of what happens at the second and third stages. The asymmetry Seymour found is therefore attributable to the difference in the interpretation times of *above* and *below* at the first stage of processing.

Seymour (1969) has recently reported a remarkable asymmetry in people's judgements of spatial location. He presented subjects with either the printed word *above* or the printed word *below* simultaneously with a configuration of a small circle either ABOVE or BELOW a large centrally located reference square. Seymour measured how long it took the subjects to judge whether or not the word correctly described the spatial location of the circle. The important result was that subjects required much less time to verify that *above* (the word) matched ABOVE (the circle's location) than that *below* matched BELOW. This finding implies that subjects do not use a simple search-and-confirm strategy of decoding the word and then looking at the critical part of the display to see if the circle is there.

Seymour attributed these results to the way the subjects scanned the display. The theory he proposed assumed that subjects first generate a model (some internal representation) of the display from the word *above* or *below* and then scan the model and display simultaneously from the top down until they locate the circle in both the model and the display. The critical part of the theory is the top–down scanning hypothesis. The assumption is that subjects invariably scan from top to bottom and, as a consequence, they are able to verify the critical upper part of the ABOVE display more quickly than the critical lower part of the BELOW display.

The purpose of the present study is to consider an alternative to the scanning hypothesis, one that is based on the abstract interpretations of the word and display. The alternative is meant to correct two weaknesses in Seymour's account. First, it seems unlikely that scanning is invariably top–down; rather, the direction of scanning will depend on specific requirements of the task. Second, Seymour's comparison operation seems too concrete for this task. His notion that subjects generate a concrete internal model from the more abstract meanings of *above* and *below* is much like Posner's (1969) concept of generation, the 'subjects' ability to go from a general code to one of greater specificity'. According to Posner, for example, subjects

can generate a visual image from the name of a letter and then use the image to compare against an actual displayed letter. But in a task like Seymour's it seems unlikely that subjects rely on this general-to-specific generation of a model, since they should be able to carry out the task even when it is impossible to scan. For example, when the top half of the display is always concealed, there is still enough information visible to decide whether *above* or *below* matches ABOVE or BELOW. But in this example, it is the interpretation of the display, not the display itself, that is required to carry out the task. That is, we suggest that subjects are more likely to generate an abstract representation out of the more complex information-rich display.

In the alternative we consider, the subjects first interpret, or represent mentally, both word and configuration as abstract symbols. The point of this process is that the subjects need to represent both the word and the display in the same format so that they can then test to see whether the two representations are the same or different. The critical feature of this account is that the mental representation of the word *above* is simpler than the word *below*, so subjects are able to encode *above* faster than *below*. The relative ease of encoding the ABOVE and BELOW displays depends critically on the visual attributes of each display. Furthermore, the encoding of the word and the encoding of the picture are independent of each other, and both are independent of the comparison process that follows. Evidence for this account will be considered more fully in the discussion that follows Expt. IV.

The present study consists of four experiments. The first is a replication of Seymour's results on the apparatus in our laboratory. The second is a repeat of the first with arrows in place of the words *above* and *below*; the aim is to demonstrate that a top-down scanning strategy is not obligatory. The third experiment, very similar to the first two, utilizes the words *present* and *absent*; the purpose here is to show a similar asymmetry between *present*–PRESENT and *absent*–ABSENT that cannot be due to a top-down scanning strategy. The last experiment again uses *above* and *below* but with displays that cannot be scanned from the top down; the purpose is to show that the same asymmetry occurs without scanning and so the *semantic* interpretations of *above* and *below* are crucial to Seymour's results. A discussion of possible explanations of these data follows the last experiment.

EXPERIMENT I

Method

The four types of displays prepared for Expt. I each consisted of a square with a smaller circle directly above or below the square; the figures were drawn with India ink and the circle was filled in to make a solid disk. At the centre of the square there was either the word *above* or the word *below* typed in Elite type. The displays, viewed through a Polymetric two-field tachistoscope at a distance of 18 in., subtended a visual angle of 4°. The four display types will be referred to as *above*–ABOVE, *below*–BELOW, *above*–BELOW, and *below*–ABOVE, depending on whether the word was *above* or *below* and whether the circle was ABOVE or BELOW the square. The first two displays – *above*–ABOVE and *below*–BELOW – will be called the 'true' conditions, since the word is a true description of the configuration; the last two displays will be called the 'false' conditions. The four displays are illustrated at the top of Table 4.

A deck of 28 stimulus cards was drawn up such that each of the four displays occurred seven times; one set of four displays was used for lead-in trials before a block and six sets were used to make the block of 24 trials. Each of 12 subjects was given four blocks of trials for a total of

96 trials plus 16 lead-in trials; the stimulus cards were in a different random order for each subject and each block of trials.

On each trial the subject was presented one of the four displays and required to push either a 'true' or a 'false' button as quickly as possible, and his latency was measured to the nearest 0·01 sec. Half the subjects pushed 'true' with their right thumb and half pushed 'true' with their left thumb. The subject was instructed to respond as quickly as possible without error. After each trial the subject was informed of his latency if he was correct, and only that an error had occurred if he was incorrect.

To initiate the next trial, the subject pushed a 'start' button (situated midway between the 'true' and 'false' buttons) and 1 sec. later – time enough for the subject to place his thumbs in the proper position – the next display appeared. The subjects were 12 Carnegie-Mellon University students fulfilling a course requirement for introductory psychology.

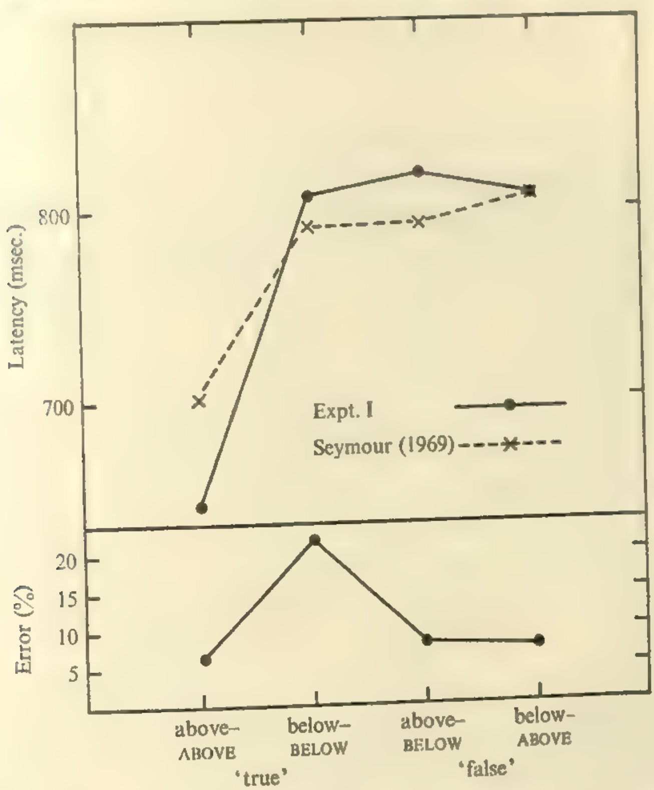


Fig. 1. Mean latencies from Expt. I and Seymour's (1969) results as a function of the words *above* and *below* and the displays ABOVE and BELOW. The means are averaged over 12 subjects, four blocks of trials and six or fewer correct responses per block.

Results

The overall mean latencies of the four conditions, shown in Fig. 1, replicated quite nicely the pattern of latencies found by Seymour, also shown in Fig. 1. The basic latency analysis was carried out on the mean latencies of the six or fewer correct responses for each condition for each block of each subject; thus there were 16 mean latencies for each of 12 subjects, or 192 means. An analysis of variance showed the following significant effects: (1) displays with *above* were verified 75 msec. faster than those with *below* ($F = 11{\cdot}2$; d.f. = 1, 11; $MS_e = 23{,}900$; $P < 0{\cdot}01$);

(2) displays with the circle ABOVE were verified 85 msec. faster than those with the circle BELOW ($F = 17{\cdot}8$; d.f. = 1, 11; $MS_e = 19{,}700$; $P < 0{\cdot}005$); (3) the 'true' displays were verified 85 msec. faster than the 'false' displays, i.e. there was a significant interaction between the *above below* variable and the ABOVE BELOW variable ($F = 15{\cdot}2$; d.f. = 1, 11; $MS_e = 23{,}200$; $P < 0{\cdot}005$); (4) subjects became faster as the experiment progressed ($F = 9{\cdot}88$; d.f. = 3, 33; $MS_e = 10{,}600$; $P < 0{\cdot}001$). No interactions between blocks and the four display conditions approached significance.

Unlike Seymour's data, the present results showed more errors in one condition than in the others. An analysis of variance showed significantly more errors on the *below*–BELOW condition than on the other three ($F = 25{\cdot}2$; d.f. = 1, 33; $MS_e = 73{\cdot}5$; $P < 0{\cdot}001$). This contrast accounted for 99 per cent of the variance between conditions and the other three conditions were not significantly different from each other.

EXPERIMENT II

The previous result, then, confirms Seymour's important finding that *above*–ABOVE takes considerably less time to verify than *below*–BELOW. If it is true, as Seymour suggests, that this result is due to an invariant top–down scanning process, then the same result should occur when *above* is replaced by the UP arrow ↑ and *below* by the DOWN arrow ↓, since the UP and DOWN arrows 'mean' the same thing in this task as *above* and *below*.

Expt. II was identical in all respects to Expt. I, except that small UP and DOWN arrows were used in place of the words *above* and *below*. The four conditions will be referred to as UP–ABOVE, DOWN–BELOW, UP–BELOW and DOWN–ABOVE, replacing *above*–ABOVE, *below*–BELOW, *above*–BELOW, and *below*–ABOVE of Expt. I, respectively. The subjects were 12 more students from the same source as those in Expt. I.

Table 1. *Mean latencies and mean percentages of errors of Expt. II*

Condition	Response	Latency (msec.)	Percentage errors
UP–ABOVE	'true'	597	4·8
DOWN–BELOW	'true'	591	5·2
UP–BELOW	'false'	615	7·1
DOWN–ABOVE	'false'	628	4·8

Results

The pattern of latencies in Expt. II was clearly symmetrical, in contrast to the previous results. The mean latencies and error percentages from Expt. II, calculated in analyses identical to those of Expt. I, are shown in Table 1. The analysis of variance produced the following results: (1) the UP displays were verified a small non-significant 4 msec. faster than the DOWN displays; (2) the ABOVE displays were verified a small non-significant 9 msec. slower than the BELOW displays; (3) the 'true' displays were verified a marginally significant 28 msec. more quickly than the 'false' displays ($F = 4{\cdot}81$; d.f. = 1, 11; $MS_e = 7620$; $P \approx 0{\cdot}05$); (4) there was a slight (30 msec.) non-significant decrease in latencies as the experiment progressed. These results are obviously quite different from those in Expt. I. This is especially true of results 1 and 2, which show negligibly small differences (4 and -9 msec.)

in this experiment in contrast to the large significant differences (75 and 85 msec.) in Expt. I. The prediction from the scanning hypothesis that UP and DOWN arrows should produce the same effects as *above* and *below* is therefore fully disconfirmed.

The percentages of errors in Expt. II also patterned quite differently from those in Expt. I. Unlike Expt. I errors, there was no significant difference among the four conditions.

Discussion

Just how did the Expt. II subjects go about verifying the displays? From the data it appears that they followed the simple strategy of (1) deciding which direction the arrow was pointing and (2) looking at the region indicated by that arrow to see if the circle was there or not. If the circle was there, they pushed the 'true' button; if not, they pushed the 'false' button. In scanning the display they apparently did not scan from the top of the display but rather looked immediately at the critical area indicated by the arrow. This simple search-and-confirm operation is just the kind that Seymour said was impossible. Since subjects did not find it necessary to scan from top to bottom in Expt. II, it is highly unlikely that they did so in Expt. I either. The explanation of the results of Expt. I must lie elsewhere, namely in the semantic interpretation of the words *above* and *below*.

EXPERIMENT III

This experiment was designed to show that the asymmetrical results of Expt. I are not unique to judgements of verticality. It can be argued that the mental operations for coding *present–absent* are similar in certain respects to those for *above–below*, and that a similar asymmetry will occur when judgements of presence or absence of the circle are made. In particular, it is predicted that the judgement *present*–PRESENT will be simpler and take less time than *absent*–ABSENT. The predicted result cannot be attributed to some top–down scanning process since information about presence or absence can occur in only one spatial location.

Method

In Expt. III, subjects were required to decide whether the word *present* or *absent* was 'true' or 'false' with respect to a circle under the four conditions, *present*–PRESENT, *absent*–ABSENT, *present*–ABSENT, and *absent*–PRESENT. In displays like those in Expt. I the word *present* or *absent* was typed within the reference square, and the circle was PRESENT or ABSENT above the square on two blocks of trials and below the square on the other two blocks. These two conditions are referred to as top-visible and bottom-visible, respectively. It should be made clear that when the circle is ABSENT the stimuli in the top-visible and bottom-visible conditions are identical – although, of course, their effect could be different by virtue of the other stimuli in these conditions.

Expt. III was identical in other respects to Expt. I, e.g. in its use of practice and experimental trials, blocks and number of trials, and so on. But to balance the effects of viewing conditions, subjects were divided into two groups. One group received the viewing conditions in the following order on the four blocks of trials: top-visible, bottom-visible, bottom-visible, top-visible. The other group received the complement of this ordering. Within each group, half the subjects used the right button for 'true' and half used the left button. The subjects were again 12 students drawn from the same source as those in Expt. I.

Results

According to the overall mean latencies and error percentages in Table 2, the two main results of Expt. III are (1) that *present*–PRESENT and *absent*–ABSENT are asymmetrical in much the same way that *above*–ABOVE and *below*–BELOW are asymmetrical, and (2) that it makes little difference whether the circle is expected above or below the square. As for the second finding, an analysis of variance of the 192 mean latencies showed that the 20 msec. difference between the top- and bottom-visible conditions was negligible ($F < 1$). The analysis showed only four significant effects: (1) the *present* displays were verified 145 msec. faster than the *absent* displays ($F = 14{\cdot}4$; d.f. = 1, 11; $MS_e = 69{,}900$; $P < 0{\cdot}005$); (2) the displays with the circle PRESENT were verified 74 msec. faster than those with the circle ABSENT ($F = 12{\cdot}1$; d.f. = 1, 11; $MS_e = 21{,}500$; $P < 0{\cdot}01$); (3) 'true' displays were verified 53 msec. faster than 'false' displays ($F = 9{\cdot}28$; d.f. = 1, 11; $MS_e = 14{,}500$; $P < 0{\cdot}025$); and (4) latencies again became shorter as the experimental session progressed ($F = 45{\cdot}2$; d.f. = 1, 11; $MS_e = 31{,}600$; $P < 0{\cdot}001$). The error percentages showed much the same pattern as the mean latencies.

Table 2. *Mean latencies and mean percentages of errors for Expt. III*

Condition	Response	Condition: Top-visible	Bottom-visible	Mean
	Latencies (msec.)			
present–PRESENT	'true'	738	755	746
absent–ABSENT	'true'	978	952	965
present–ABSENT	'false'	846	900	873
absent–PRESENT	'false'	927	961	944
Average		872	892	882
	Error percentages			
present–PRESENT	'true'	2·1	4·2	3·1
absent–ABSENT	'true'	16·0	15·3	15·6
present–ABSENT	'false'	4·9	8·3	6·6
absent–PRESENT	'false'	4·9	6·3	5·6
Average		7·0	8·5	7·7

Discussion

Expt. III, then, demonstrates that *present* and *absent* are asymmetrical just as *above* and *below* were, in spite of the fact that *present* and *absent* require no scanning. The information about the presence or absence of the circle was available at only one spatial location on any one block of trials. Furthermore, since the latencies were virtually identical for the top- and bottom-visible conditions, the conclusion is strengthened that subjects do not invariably scan displays from top down. The *present–absent* results, then, must be accounted for by underlying mental representations, not by a scanning mechanism, and evidence is mounting that the same is true for *above–below*.

Experiment IV

One might argue that Expt. II was a 'purely' perceptual task, containing, as it did, only arrows, squares and circles, and that because no linguistic interpretation of the words *above* and *below* was needed, the task somehow allowed subjects to bypass the requirement that they scan configurations from top to bottom. Posner & Mitchell (1967), for example, were able to identify different levels of processing involved in match mismatch judgements for letters of the alphabet; the comparison could be carried out on the physical forms of the letters, on the letter names, or on some classification rule (consonant–vowel), depending upon the depth of processing required. One might argue, then, that the words *above* and *below* themselves constrain the subjects to scan configurations as Seymour proposed. This is not unreasonable since *above* and *below* both refer to verticality and hence they might force subjects, in some unknown way, to scan from the top down. Note, however, that this very restricted hypothesis is not the one proposed by Seymour, since he appears to posit the scanning strategy no matter what the task. Expt. IV was designed to eliminate even this last thread of hope for the scanning hypothesis.

The strategy of Expt. IV – similar to that of Expt. III – was to give subjects displays which they could not verify by scanning; instead they were always forced to inspect only one part of the display and make inferences about another part. If the *above*–ABOVE *below*–BELOW asymmetry still persists with these displays, the scanning hypothesis can be ruled out.

Method

Expt. I was repeated, but the spatial location either above or below the square was masked. Thus when the top-visible mask was used, the subjects could see the circle when it was above the square, but nothing when it was below the square; in the latter case, they had to infer that the circle was below the square by noting that the circle was not above the square. The analogous situation held, of course, with the bottom-visible mask. In Expt. IV, then, top–down and search-and-confirm strategies were of no use since within any block of trials only one part of the display contained information.

Expt. IV was identical in all respects to Expt. III, except that Expt. IV used the same four displays as Expt. I. Again top-visible and bottom-visible conditions were given to each subject in a counterbalanced ABBA design. The subjects were 12 students drawn from the same source as those in Expts. I, II and III.

Results

The important finding of Expt. IV was that the pattern of results for the top-visible condition closely paralleled that of Expt. I, but the pattern was quite different for the bottom-visible condition. The overall mean latencies for the eight conditions of Expt. IV are shown in Fig. 2 along with the mean latencies for Expt. I, labelled 'both-visible'. The comparison of Expts. I and IV can be seen more clearly in Fig. 3, which shows differences between the both-visible condition (Expt. I) and the top-visible and bottom-visible conditions. A preliminary analysis showed no differences between the top-visible and both-visible conditions; the 94 msec. overall difference between groups was not significant. A comparison between bottom-visible and both-visible conditions, however, showed a significant overall difference of 158 msec. between the groups ($F = 9{\cdot}08$; d.f. $= 1, 22$; $MS_e = 65{,}530$; $P < 0{\cdot}01$).

More importantly, there was a strong interaction of viewing conditions with the ABOVE-BELOW variable ($F = 17{\cdot}1$; d.f. = 1, 22; $MS_e = 6251$; $P < 0{\cdot}001$). This result can be more easily seen in Fig. 3, where two conditions are seen to be considerably slower than the other six. These are the two conditions in which the subject had to infer that the circle was ABOVE. Whereas this inference is apparently quite difficult, the inference that the circle is BELOW requires no more time in the top-visible condition than in the both-visible condition.

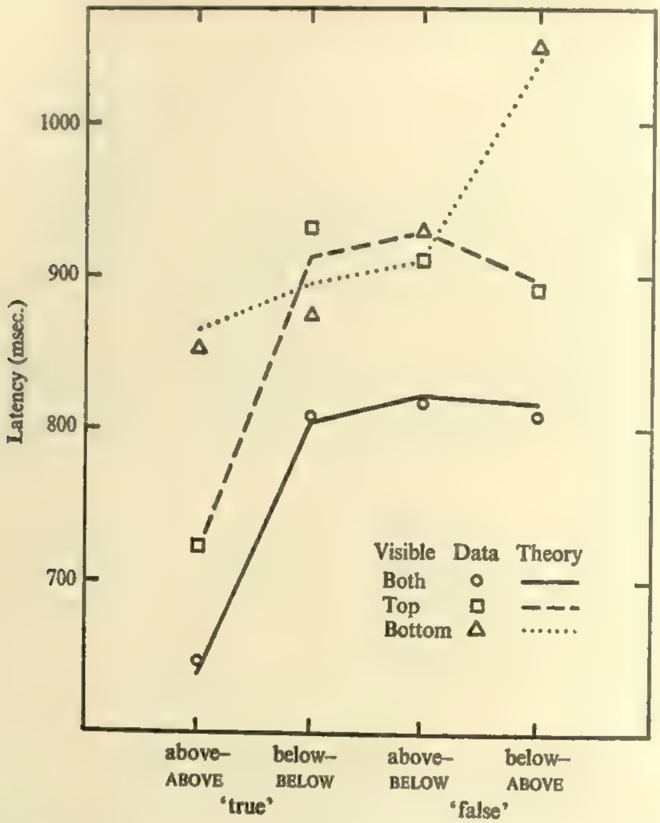


Fig. 2. Mean latencies for Expt. I, labelled 'both-visible', and the top- and bottom-visible conditions of Expt. IV. Also shown are the predicted latencies derived from the model of Clark & Chase (in preparation).

The principal difference among the both-, top- and bottom-visible conditions was to be found in the differences between the ABOVE and BELOW displays. The differences between *above* and *below* displays, and between 'true' and 'false' displays, remained unaffected by the visibility conditions. The analysis of variance for Expt. IV showed the following significant results. (1) The displays with *above* were verified 83 msec. faster, on the average, than those with *below* ($F = 8{\cdot}56$; d.f. = 1, 11; $MS_e = 39{,}500$; $P < 0{\cdot}025$); this difference is close to the 75 msec. found in the both-visible condition of Expt. I. Furthermore, the interaction between the *above–below* factor and the top- *v.* bottom-visible factor did not approach significance. (2) The 'true' displays were verified 100 msec. faster, on the average, than the 'false' displays ($F = 19{\cdot}6$;

d.f. = 1, 11; MS_e = 25,000; $P < 0{\cdot}001$); this difference is close to the 85 msec. found in the both-visible condition. Again, there was no interaction of 'true' 'false' with the visibility conditions. (3) There was a strong interaction between ABOVE BELOW and the viewing conditions such that the ABOVE displays were verified 115 msec. faster in the top-visible condition, but 49 msec. slower in the bottom-visible condition ($F = 23{\cdot}8$; d.f. = 1, 11; MS_e = 13,400; $P < 0{\cdot}001$). The ABOVE displays were verified 86 msec. faster than the BELOW displays in the both-visible condition. (4) The overall mean for the top-visible display was 63 msec. faster than

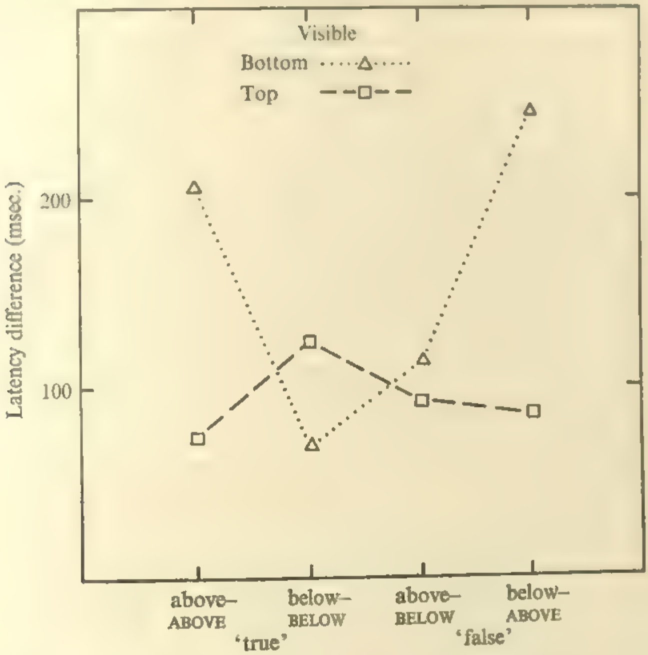


Fig. 3. The difference between the mean latencies of top- and bottom-visible conditions and that of the both-visible condition.

Table 3. *Mean percentages of errors in Expt. IV*

Condition	Response	Condition: Top-visible	Bottom-visible	Average
above–ABOVE	'true'	13·2	7·6	10·4
below–BELOW	'true'	16·0	13·2	14·6
above–BELOW	'false'	2·1	5·6	3·8
below–ABOVE	'false'	4·9	11·8	8·4
Means		9·0	9·6	9·3

the mean of the bottom-visible display ($F = 7{\cdot}45$; d.f. = 1, 11; MS_e = 25,900; $P < 0{\cdot}025$). (5) Finally, there was again an overall decrease in latencies with practice ($F = 155$; d.f. = 1, 11; MS_e = 53,900; $P < 0{\cdot}001$). The error percentages were more variable than the mean latencies, but they again showed the highest error rates in the *below*–BELOW condition (Table 3).

Discussion

The significance of Expt. IV is (1) that it suggests that abstract mental representations are used in making spatial judgements and (2) that it eliminates all hope for the scanning hypothesis.

Probably the most important finding was that the *above* displays were about 80 msec. faster than the *below* displays no matter whether the whole display or just the top or bottom was visible. Similarly, the 'true' displays were verified about 90 msec. faster than the 'false' displays regardless of the three visibility conditions. The only latencies affected by the three visibility conditions were the differences between the ABOVE and BELOW displays. This finding suggests that the mental processes that result in the *above–below* and 'true'–'false' differences can be separated out from those that result in the ABOVE–BELOW difference. More will be said about this later.

A subsidiary finding was that the pattern of latencies in the top-visible condition was approximately the same as in the both-visible condition, but the pattern in the bottom-visible condition was considerably different from the other two. This finding suggests that subjects employ approximately the same processes in the both- and top-visible conditions, but quite different processes in the bottom-visible condition. Put more specifically, the process of seeing that a circle is BELOW and the process of inferring that a circle is BELOW must be quite similar ones, but the process of seeing that a circle is ABOVE and the process of inferring that a circle is ABOVE must be quite different from each other. More will be said later about this too.

As for the scanning hypothesis, the results of Expt. IV finally rule out the hypothesis completely. According to the subsidiary finding, the both- and top-visible conditions resulted in the same patterns of latencies in spite of the fact that a scanning strategy was impossible in the top-visible condition. Therefore, since the *above*–ABOVE *below*–BELOW asymmetry can occur in the absence of scanning, the asymmetry cannot be accounted for by the scanning hypothesis.

The results also rule out the possibility of explaining the *above*–ABOVE *below*–BELOW asymmetry solely by an attention hypothesis. According to this view, people normally attend to the top of the picture, and the BELOW displays are processed by inference. Thus, *above*–ABOVE is faster than *below*–BELOW because, in the *below*–BELOW case, people have to make a time-consuming inference that the circle is BELOW; as a consequence, masking the bottom of the displays should have no effect since people infer that the circle is BELOW anyway. This explanation fails, however, to account for the bottom-visible latencies. Under this assumption, the bottom-visible latencies should show the same pattern as the top-visible latencies, once *above* and *below*, and ABOVE and BELOW, have been interchanged, because now people attend directly to the BELOW displays but make the time-consuming inference for the ABOVE displays. But the pattern in the results is nothing like this expected pattern. To take an example, in the bottom-visible condition, *below*–BELOW should be faster than *above*–ABOVE, since *below*–BELOW can be verified directly, whereas *above*–ABOVE requires an inference. Furthermore, this *below*–BELOW *above*–ABOVE difference should be the same size as the *above*–ABOVE *below*–BELOW difference in the top-visible condition. But in fact, in the bottom-visible condition, *above*–ABOVE was

still faster than *below*–BELOW, completely contrary to this expectation. Results like these, then, simply cannot be accounted for by an attentional-scanning strategy.

The findings of Expt. IV therefore suggest that another type of theory altogether is required to account for the verification process in Expts. I and IV. We will now present a theory that begins with abstract interpretations of *above*, *below*, ABOVE, and BELOW. It is based on a considerable amount of subsequent work on the *above–below* asymmetry; the present results are clearly too meagre to support an elaborate model by themselves. We will describe the parts of the theory that apply to the present experiments; the full theory and its justification can be found in Clark & Chase (in preparation).

A THEORY OF VERIFICATION OF SPATIAL LOCATION

The proposed explanation for Expts. I and IV is based on the premise that the time between stimulus and response is composed of a series of additive stages in the sense described by Sternberg (1969). In the first of four gross additive stages, the subject encodes the word *above* or *below*; in the second, he encodes the location of the circle; in the third, he compares his mental representations of the word and location; and in the fourth, he responds. The explanation of the *above–below* asymmetry lies completely in the representations that *above* and *below* take in the first stage.

At the first stage, the Word Encoding stage, it is assumed that words or sentences are encoded approximately in terms of their deep structure propositions (cf. Clark, 1969), which are represented as sets of symbols and rules. In this experiment we assumed that *above* and *below* actually stand for the full sentences *The circle is above the square* and *The circle is below the square*, respectively. These two sentences, in turn, contain very simple deep structure propositions that may be symbolized as (*circle above square*) and (*circle below square*), respectively. But these representations are not detailed enough, for *above* and *below* have even more abstract representations. Linguistically, *above* and *below* are closely related. Both *above* and *below* specify vertical comparison, although *above* specifies comparison in one direction and *below* specifies comparison in the opposite direction. Thus, *above* and *below* could be specified in a featural notation in which all features but one are identical for the two, and the one feature that differs specifies direction of comparison. The solution, however, is not that simple, for the word *above* is the normal or simpler way to encode two vertically arranged objects, while *below* is the more complex. To understand this, we must consider two linguistic facts about English: (1) that *A is above B* and *B is below A* are not synonymous, and (2) that verticality is always described in English in an asymmetrical fashion.

The first linguistic fact is that *A is above B* and *B is below A*, although they appear to be synonymous, differ at a critical point. The first sentence describes the position of *A* with respect to the position of *B*, whereas the second one does just the opposite. *A is above B*, for example, presupposes that the position of *B* is already known to the listener, and it asserts that *A* has a particular position with respect to *B*. Thus it is unacceptable to answer the question *Where is B?* with the answer *A is above B* (with normal intonation), since the presupposition of the answer (that the position

of B is known, conflicts directly with the presupposition of the question, that the position of B is *not* known. Therefore, *A is above B* is used to describe a picture in which B is the point of reference, whereas *B is below A* is used where A is the point of reference.

The second linguistic fact is that English normally describes verticality such that the point of reference is at the bottom of the described dimension. The only English adjectives used exclusively for describing vertical position are *high*, *low*, *tall* and *short*. It is already well known (cf. Clark, 1969) that *high* and *low* presuppose the semantically prior dimension of *height* (not *lowness*), and that they mean, roughly, 'of much height' and 'of little height', respectively. But *height* is always measured upward from a point of reference at the bottom, no matter whether much or little height is being measured. The analogous situation holds for *tall* and *short* and their semantically underlying dimension *height* (tallness). So in spite of their other differences, *high*, *low*, *tall* and *short* all presuppose a point of reference at the bottom of what is being measured or described.

The sentence *A is above B* also has the property that the point of reference, B, is at the bottom of what is being described; *B is below A* does not have this property. This suggests that the sentence *A is above B* is the normal way to describe a picture of an A above a B. Since verticality automatically brings with it the assumption that the point of reference is at the bottom, we have hypothesized that the semantic representation of *above* is formed in one step (since it conforms to this assumption), whereas that of *below* requires one additional step specifically to alter this automatic assumption (cf. Clark & Chase, in preparation, for details). In short, the *above* and *below* in (*circle above square*) and (*circle below square*) should be thought of as amalgamations of more primitive features, with the representation of *below* requiring one more step to form than the representation of *above*; likewise, *circle* and *square* represent more detailed semantic specifications of the abstract concepts of 'circle' and 'square'.

These considerations have one important consequence on the Word Encoding stage. Since *above* is represented more simply than *below*, the *above* displays should be represented more quickly than the *below* displays. The amount of time by which *above* is represented faster than *below* is given by the parameter a in this model.

At the second stage, the Picture Encoding stage, it is assumed that the ABOVE and BELOW locations of the circle are encoded, respectively, as (*circle above square*) and (*circle below square*). But the three visibility conditions vary in how easy it is for the subject to set up the two representations. Apparently, subjects normally attend to the upper location of the circle, hence they normally (or often) only *infer* (*circle below square*) in the BELOW displays. When either the bottom or top of the display is hidden from view, of course, the subject has to infer the hidden location. In short, (*circle above square*) should be encoded some amount b_1 faster than (*circle below square*) in the both-visible condition, some amount b_2 faster in the top-visible condition, but some amount b_3 slower in the bottom-visible condition.

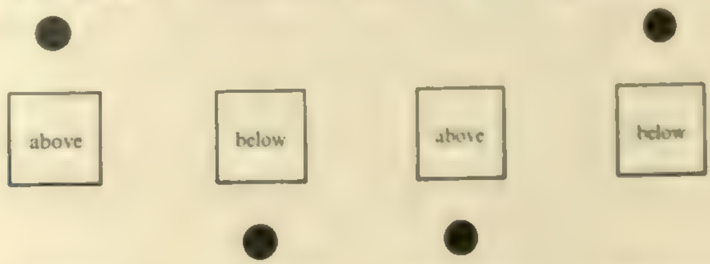
At the third stage, the Comparison stage, it is assumed that the subject (in these experiments at least) must carry out two time-consuming mental operations. The purpose of these operations is to keep track of an indexical 'truth value' of the word *vis-à-vis* the picture and to produce its correct value in the end as *true* or *false*

i.e. match or mismatch. The truth value is assumed to be *true* at the beginning, and the finding of a mismatch between the word and picture representations changes that *true* to *false*. Thus Operation 1 compares the representation of the word with that of the picture. If the two representations match completely, then the presupposed truth value *true* is retained. If the two representations do not match, Operation 1*a* changes the truth value from *true* to *false*.

The consequence of Operations 1 and 1*a* of the Comparison stage is to cause the two 'false' conditions to be slower than the two 'true' conditions. The first time-consuming operation (Operation 1) is required by all four displays – *above* ABOVE, *below* BELOW, *above* BELOW and *below* ABOVE – so it adds an equal increment of time to each of these conditions. But the second operation (Operation 1*a*) is required only of the two 'false' displays, *above* BELOW and *below* ABOVE, so it adds an increment only to these two displays. This increment of time is given in the model by the parameter c. Thus this model of the Comparison stage predicts that the two 'true' displays should be verified a fixed time increment c faster than the two 'false' displays.

Table 4. *The four displays and the theoretical comparison process for each*

The displays



Description of display	above–ABOVE	below–BELOW	above–BELOW	below–ABOVE
Stages				
Word encoding	(cir above sq)	(cir below sq)	(cir above sq)	(cir below sq)
Picture encoding	(cir above sq)	(cir below sq)	(cir below sq)	(cir above sq)
Comparison				
Initial truth value	true	true	true	true
Operation 1	match	match	mismatch	mismatch
Operation 1*a*	—	—	true → false	true → false
Response	'true'	'true'	'false'	'false'

Finally, it is assumed that the Response stage merely takes the final truth value of the Comparison stage – *true* or *false* – and converts it into a push of the correct 'true' or 'false' button. This process has no differential consequences for the latencies of the four displays. The four stages are shown in abbreviated form for the four displays in Table 4.

The proposed model, then, makes use of five parameters (a, b_1, b_2, b_3 and c) to predict the differences in latencies among the four displays within the three visibility conditions. From the appropriate subtractions in the data at the top of Table 5, a is estimated to be 81 msec.; b_1, 86 msec.; b_2, 115 msec.; b_3, 49 msec.; and c, 96 msec. These parameters, along with the 'base' times, t_1, t_2 and t_3, can be combined appro-

priately, as shown in the middle of Table 5, to form the predicted times at the bottom of Table 5. The predicted latencies are very close to the actual latencies, as can be seen by referring back to the two sets of times in Fig. 2. To evaluate the theory, the 12 d.f. associated with the data of Table 5 are partitioned into the 3 d.f. for estimating the 'base' times and the 9 d.f. within viewing conditions. Of the 9 d.f. within viewing conditions, the model uses 5 d.f. to estimate its parameters, and 4 d.f. are left for residual variance or lack of fit. The model accounts for over 98 per cent of the variance among the 9 d.f. within conditions; the residual variance does not approach significance when tested against the pooled error term. The lack of fit is best evaluated by calculating the root mean squared deviation (RMSD) and then comparing this RMSD with the smallest parameter estimate of the model (Sternberg, 1969). In this case, the RMSD of 13 msec. with 4 d.f. indicates a good fit. Thus the present model is accurate in predicting the verification latencies in Expts. I and IV.

Table 5. *Actual latencies, theoretical model and predicted latencies for Expts. I and IV*

	Visibility condition		
Display	Both-visible	Top-visible	Bottom-visible
Actual latencies			
above–ABOVE	647	722	852
below–BELOW	807	932	875
above–BELOW	818	911	931
below–ABOVE	807	892	1052
Theoretical model			
above–ABOVE	t_1	t_2	t_3+b_3
below–BELOW	t_1+a+b_1	t_2+a+b_2	t_3+a
above–BELOW	t_1+b_1+c	t_2+b_2+c	t_3+c
below–ABOVE	t_1+a+c	t_2+a+c	t_3+a+b_3+c
Predicted latencies			
above–ABOVE	639	718	864
below–BELOW	805	913	896
above–BELOW	821	929	911
below–ABOVE	816	895	1041

It would have been more elegant if b_1, b_2 and b_3 had been equal, indicating that the subjects were simply inferring BELOW in the both- and top-visible conditions and ABOVE in the bottom-visible condition and that this inference took the same amount of time in each instance. Under the assumption that the b's are equal, in fact, the model (b = 84 msec.) still accounts for 95 per cent of the variance with only 3 d.f., and the RMSD of 23 msec. with 6 d.f. might be considered a marginal fit. However, there are two points which deviate (significantly) by about 50 msec. from the model, and it can only be concluded that the problem is due to different Picture Encoding times (b) for the different viewing conditions. The reason the b's vary is certainly not clear, although there are many picture encoding strategies that might account for the variation. In the subsequent experiments in which more care was taken to counterbalance for certain kinds of picture encoding strategies, the b effect disappeared altogether (Clark & Chase, in preparation). So the b effect appears to be quite specific to the several viewing conditions of this experiment, whereas the a and c effects are quite independent of the viewing conditions.

The estimates for a, b_1 and c for Seymour's data are 53, 38 and 53, respectively. These estimates are considerably lower than the comparable estimates from Expt I alone of 75, 86 and 86, respectively. It is not clear why these two sets of parameters should be so different.

The most important thing this model has accomplished has been to separate the two encoding difficulties from each other and from the verification difficulties. First of all, this model asserts – in agreement with the results of Expts. I and IV – that *above* takes 81 msec. less time to interpret than *below*, no matter what the display is and no matter what comparison has to be made. This assumption is completely in keeping with all the subsequent data of Clark & Chase (in preparation), which show (1) that this *above–below* difference is consistently about 80 to 100 msec. and (2) that this difference is independent of the display to be verified as well as other attributes of the task. Secondly, the model also asserts that the 'true'–'false' difference is also constant at 96 msec. and is independent of specific attributes of the displays. Again, this is quite consistent with all the results of Clark & Chase (in preparation). But most importantly, the model quite appropriately asserts that the three visibility conditions affect only one thing – the relative difficulty of encoding the ABOVE and BELOW positions of the circle. In other words, the interpretations of *above* and *below* are not affected by variations in the viewing conditions, nor is the verification process itself affected: the only part affected by covering up various parts of the display is the process of representing the location of the circle in memory, just as one would expect. In short, the present model has several advantages over Seymour's attentional-scanning hypothesis. First, it correctly accounts for the data in Expts. I and IV. Secondly, the model explicitly separates out the subprocesses of the task and correctly correlates changes in the experiment with the appropriate changes in the model. And thirdly, the present model has wide generality, for in its more general form it is able to account for many other locational and non-locational tasks for which the attentional-scanning hypothesis makes either incorrect predictions or no predictions at all.

At this point, one might still suspect that the present results can somehow be accommodated by a mental imagery model (e.g. DeSoto *et al.*, 1965; Huttenlocher, 1968). However, such a model fails to explain our later results (Clark & Chase, in preparation), and on quite independent grounds. The problem is with negative sentences such as *A isn't above B*.

First, it is impossible in general to form a single image, or even an image surrogate as in Seymour's proposal, of negative sentences (e.g. *A isn't above B*). The general property of such sentences is that they indicate what a picture is not, and only rarely does this uniquely define what a picture is. But since the imagery model requires that a single image be matched against the picture, the model cannot work in general for negative sentences. And since it cannot account for negatives, it is unlikely that it accounts for positive sentences either. Rather, we suppose that both positive and negative sentences are verified in the same way, and this assumption allows the imagery model to be rejected on *a priori* grounds.

The imagery model, however, can be ruled out on empirical grounds as well. To apply the imagery model to the Clark & Chase results, we would have to make two assumptions: (1) that the imagery model could provide a plausible account of why

an image of *A above* is easier to construct than an image of *B below* for positive sentences, and (2) that subjects use the same procedure to form a single image from negative sentences like *B isn't above A* and *A isn't below B*, a strategy that was possible in this particular experiment. The results were incompatible with the imagery model on two counts. First, it was found that *B isn't above A* took less time to verify than *A isn't below B*. By the imagery model, forming an image from bottom up would have to be easier than forming an image from top down for these negative sentences, whereas exactly the opposite would hold true for positive sentences (cf. Assumption 1). Hence the imagery assumption is not consistent with these results. Secondly, it was found that for positive sentences true responses are faster than false responses, but for negative sentences true responses are *slower* than false responses. To account for this result, the imagery model would have to assume that each mental image contains information about whether it was constructed from a positive or negative sentence. Images as normally conceived simply do not do this. In contrast, the linguistic model we have proposed correctly accounts for these results in every detail.

In summary, the present experiments have allowed us to make both a negative and a positive point. First, Expts. II, III and IV rule out an attentional-scanning process as a possible explanation of the *above*–ABOVE *below*–BELOW asymmetry in Expt. I. Secondly, these same experiments suggest a model that consists of at least four additive stages: a Sentence Encoding and a Picture Encoding stage, in which the word and picture are represented mentally as abstract symbols, a Comparison stage, in which the resulting two representations are compared symbolically, and a Response stage where the result of the Comparison – either *true* or *false* – is mapped on to a 'true' or 'false' button press. And it was shown that it is the semantic interpretation of the words and pictures that lies at the heart of the *above*–ABOVE *below*–BELOW asymmetry.

This work was supported by Public Health Service Research Grant MH-07722 from the National Institute of Mental Health. We would like to thank Calvin Nodine for suggesting Expt. II and Michelene Chi Chase for running subjects. We are indebted to Michael Posner for his critical comment on an earlier draft.

REFERENCES

CLARK, H. H. (1969). Linguistic processes in deductive reasoning. *Psychol. Rev.* **76**, 387–404.

DESOTO, C., LONDON, M. & HANDEL, S. (1965). Social reasoning and spatial paralogic. *J. Person. soc. Psychol.* **2**, 513–521.

HUTTENLOCHER, J. (1968). Constructing spatial images: a strategy in reasoning. *Psychol. Rev.* **75**, 550–560.

POSNER, M. I. (1969). Abstraction and the process of recognition. In G. H. Bower & J. T. Spence (eds.). *Advances in Learning and Motivation*, vol. 3. New York: Academic Press.

POSNER, M. I. & MITCHELL, R. F. (1967). Chronometric analysis of classification. *Psychol. Rev.* **74**, 392–409.

SEYMOUR, P. H. K. (1969). Response latencies in judgements of spatial location. *Br. J. Psychol.* **60**, 31–39.

STERNBERG, S. (1969). The discovery of processing stages: extensions of Donders' method. *Acta psychol.* **30**, 276–315.

(*Manuscript received* 19 *June* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 327–334
Printed in Great Britain

SOME FURTHER OBSERVATIONS ON THE MEASUREMENT OF CLUSTERING IN FREE RECALL

By E. C. DALRYMPLE-ALFORD*

Department of Psychology, American University of Beirut, Lebanon

Properties of recently proposed measures of clustering in recall are considered. Applying some measures to computer generated sequences revealed that some of them are liable to give a spurious picture of differences in the tendency for items to cluster if sequences they also differ in length.

The study of organizational factors in memory has long since progressed beyond the stage of demonstrating that such factors may play an important role in the retention and retrieval of information. Unfortunately, there is reason to suspect that we may be building up a somewhat spurious picture of some aspects of the phenomena simply because we are using measures whose properties have been insufficiently explored. This seems particularly true of studies of clustering in free recall where we are faced with a proliferation of indices and an increasing feeling of discomfort that we may be dealing with artifacts of the measures adopted. The comments that follow are offered in the hope that even if the question of the most suitable measure is not entirely resolved, the dimensions of the problem will be made more clear.

Clustering in free recall is usually measured in terms of the number of 'repetitions' that occur in the subject's output, a 'repetition' being the occurrence of an item immediately after another item of the same type (e.g. semantic category). Since repetitions may occur within a series even when the series is ordered randomly, a satisfactory measure of clustering would indicate zero genuine clustering when the observed number of repetitions is no different from what would be expected if the series were randomly ordered. Further, a suitable measure of clustering would covary in a systematic manner with the tendency for genuine clustering to occur. Finally, since, given the same tendency for clustering, the absolute number of repetitions would increase with the length of the series, a suitable measure of clustering would be unaffected by the length of the series as long as this tendency is constant. These then are the criteria against which proposed measures of clustering should be evaluated.

Some indices of clustering

Shuell (1969), in a useful review of organizational factors and recall, has examined some of the indices that have been proposed from time to time and drawn attention to what he considers to be their limitations. This paper will not attempt to cover the same ground but will concern itself only with the ratio of repetition (RR) index (Bousfield, 1953) and some more recent proposals that his review did not deal with.

The RR index, by far the most popular, is simply the ratio of the number of repetitions (R) in a series to one less than the total number of items (N) in that

* Present address: Trent University, Peterborough, Ontario, Canada.

series. (Bousfield's original choice of denominator was N; the change to $N-1$ was probably based on the mistaken notion that the maximum possible value of R is $N-1$, when in fact it is $N-k$, where k is the number of different categories of items in the series.) The upper limit of RR is thus $(N-k)/(N-1)$, and will be 1 only when all the items in the series under consideration belong to the same category. Thus in the series AAAABB, the items (letters) are clustered maximally but the value of RR is 0·80.

When the probability of a repetition occurring is constant over all positions in the series, the RR index could appear to be the most appropriate measure of clustering for it may be shown that $R/(N-1)$ is the best estimate of the probability of a repetition occurring. However, the probability of a repetition occurring is not the best indicator of clustering. For example, a probability of 0·25 of a repetition occurring represents chance clustering when only four types of items are possible, below chance clustering when only two types may occur, but above chance clustering when more than four types may occur.

Replacing $(N-1)$ in RR by $(N-k)$ gives us the MRR index devised by A. M. Lesgold and D. Tieman which appears in Bower *et al.* (1969). With this modification we have an index whose value will be 1 when all the items in the series are clustered maximally. However, the value of this index is not necessarily 0 when the number of repetitions in the series is the minimum that is possible. For example, if the sequence consists of six As and two Bs, it is not possible to have fewer than three repetitions even if we deliberately attempt to keep letters of the same sort separated. Hence the value of MRR cannot be less than 0·5. Thus one runs the risk of concluding there is as much evidence of clustering in the sequence ABABAAAA as there is in the sequence AABBABBA. In both the maximum possible number of repetitions is 6, the observed number of repetitions is 3, and therefore the value of MRR is 0·5 for both, even though the number of repetitions in the first sequence is the minimum possible, while in the second sequence the observed number of repetitions is higher than the minimum possible, which is zero for that sequence. It is true of course that while there is the possibility of being misled by the values of this index, the risk is fairly low since the minimum number of repetitions will be greater than zero only when there is fairly extreme imbalance in the number of items in the sequence that belong to the different categories, i.e. when the number of items belonging to any category is greater than one more than half the total number of items in the sequence.

The C index, first used by Dalrymple-Alford & Aamiry (1969), takes the modification of the RR index one step further. By defining $C = (R - \min R)/(\max R - \min R)$, where $\max R$ and $\min R$ are the maximum and minimum possible values of R, respectively, we have an index that is 1 when clustering is at the maximum possible, and 0 when it is at the minimum, regardless of how many items there are in the sequence, or of their distribution among the categories involved. It should be clear that $\max R$ and $\min R$ refer to the maximum and minimum number of repetitions that could be obtained given that the subject produced the items that he did. We are concerned with clustering manifest in the output. Measures of clustering that take into consideration features of the input are very likely to involve assumptions that we may wish to test when we measure clustering.

It is worth noting that if interest lies solely in determining whether clustering

manifest in a series reflects genuine clustering, it is sufficient simply to count the number of repetitions in the series and compare this with what would be expected by chance. It is only in situations where we wish to make comparisons across conditions that an index independent of the number of items in the series, and their distribution into the different categories, is useful. To say that such an index would be useful is not to say that there are no further difficulties. The C index, for example, is independent of how many words are recalled from each category, but the expected (chance) value of C is not. At the risk of being tedious let me present the problem graphically. Consider the sequences (i) AAAABBABA, (ii) AABBBABAA, (iii) AABBCCCAB. Fig. 1 (a) represents the range of possible values of R for each case, together with observed and expected values of R. Fig. 1 (b) represents the same thing but in terms of the C index, and shows that C measures clustering on the same 'scale'

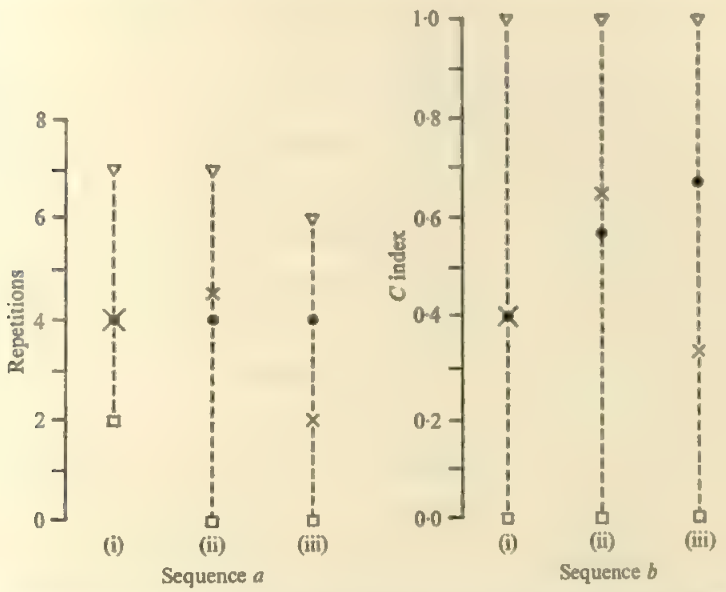


Fig. 1. Illustration of observed, expected (chance), maximum and minimum possible number of repetitions, and the C equivalents of these, for three series given in the text. ▽, maximum value; ●, observed value; ×, chance value; □, minimum value.

regardless of the range of possible values of R. But comparisons across series must also take into account the fact that the observed value of series (i) is at chance level, for series (ii) below chance, while in series (iii) it is clearly above chance. I have suggested before (Dalrymple-Alford, 1970) that comparisons be made in terms of the difference (observed C − expected C). Thus instead of comparing differences between the observed and expected number of repetitions, as Bousfield & Bousfield (1966) suggest, the recommendation is to make these comparisons taking into account the range of possible values of R. It seems appropriate to measure observed and expected clustering on a 'scale' that is invariant with the number of different types of items recalled, and in like manner to make comparisons on the same scale too.

Hudson & Dunn (1969) have recommended a different alternative to the Bousfield

& Bousfield (1966) proposal, namely that comparisons should be made in terms of the ratio of the difference between observed and expected values of R to the standard deviation of R given that the series is randomly ordered. (The standard deviation of R, on the assumption of random arrangement of the items, was derived by Barton & David, 1957, and is given more simply in the Hudson–Dunn paper.) The Hudson–Dunn measure is useful when one is faced with a single series and wishes to decide whether the clustering manifest in it is enough to warrant the conclusion that there is genuine organization of the elements on the basis of category membership or whether the clustering observed is a chance thing. Barton & David (1957) have shown that on the null hypothesis that the series is randomly ordered, the distribution of R is reasonably approximated by the normal distribution with mean and variance as given in Hudson & Dunn (1969), especially when all categories are equally represented in the series and the series contains 12 or more items. One may therefore regard the Hudson–Dunn measure as a z score to test whether the manifest clustering in a single series is significantly different from chance.

But what of comparisons involving different series? Hudson & Dunn say that their measure 'of course, produces approximately normally distributed standard scores which allow direct comparison of subjects who recall words from different numbers of categories and have different recall totals' (p. 10). But this is not so. It is only on the null hypothesis (i.e. that the arrangement of the items is random) that we can accept that the observed value of R comes from a distribution with the mean and variance that are used in the Hudson–Dunn measure. Furthermore, the normal approximation has been demonstrated only for the case when the items are randomly arranged. Thus the property that Hudson & Dunn find useful in their measure exists precisely in those circumstances where comparison becomes pointless – where clustering does not differ from what would be expected to arise by chance alone.

More recently, Dunn (1969) has proposed another measure which he called D_X. He makes the interesting point that information about categorical associations is contained in the number of items recalled from each of the categories present in the input, and then derives his measure which is similar to the Hudson–Dunn measure, except that the mean and variance are made conditional on both the input (i.e. how many items of each type were presented to the subject), and the output of the subject. As with the Hudson–Dunn measure, and therefore subject to the same criticisms, the mean and variance are determined on the assumption that the ordering of the elements in the output does not reflect genuine clustering.

One gets the impression that Dunn has misunderstood the requirements of a measure of clustering. Presumably one of our concerns is with the relationship between clustering and recall performance, e.g. is clustering associated with higher recall? (e.g. Cofer, 1967). What we seek therefore is some suitable way of measuring clustering that enables us to examine the organization of the output independently of the sheer number of items in the output, or for that matter in the input. Dunn's measure fails simply on the grounds that it is based on assumptions which we may be interested in testing when we seek a measure of clustering. The inappropriateness of Dunn's measure is also fairly clear when we consider some of the sequences Dunn uses for illustration: (i) AAAA, (ii) AAAAAAAA, (iii) AABBCCDD. The values of D_X for these are 3·278, 5·019 and 2·241 respectively. Thus we are asked to conclude that

relative to chance, clustering is greatest in (ii) and least in (iii). However there is really no way of distinguishing between perfect clustering and completely random arrangement when all the items recalled come from the same category. Since the extent of clustering present in such sequences is indeterminate, they should be eliminated from analysis. (This poses no great problem since in practice such series are encountered very rarely.) Dunn, however, disagrees. He says "in fact we must argue that an extreme amount of association has occurred with category A at the expense of all other categories" (p. 1417). This is an interesting hypothesis, and it awaits testing—but surely not with a measure that is based on the assumption that it is true.

An empirical comparison of some measures

Since demonstrations are generally more convincing than formal argument, it is interesting to see how some of the measures fare in a situation where genuine clustering is known to exist at different levels. Actual data from free recall studies are useless for this purpose since we have no means of being sure of the level of clustering without recourse to one or other of the measures. What was done therefore was to generate strings in such a way that one could know on *a priori* grounds what differences existed between sets of strings in the tendency for genuine clustering to occur. Ten sets, each consisting of 300 sequences, were generated by computer using only the letters A to E. The sets differed in the probability with which a letter was repeated at any point in a sequence. This probability, p, was varied from 0 to 0·9 in steps of 0·1. Within any set the probability that letter i would be followed by letter j was set at $\frac{1}{4}(1-p)$ for all $i \neq j$. Thus when $p = 0{\cdot}2$ all letters had an equal likelihood of occurring at any point in the sequence regardless of what had occurred immediately before: in other words, the tendency for genuine clustering was zero. With $p > 0{\cdot}2$, the tendency for clustering was above chance, while $p < 0{\cdot}2$ gave rise to below chance clustering. Within each set there were 100 sequences of length 10 items, 100 of length 20 items, and 100 of length 30 items.

The same program that generated the sequences also computed the following measures on each:

(i) $D_B = R - E(R)$,

(ii) $D_R = [R - E(R)]/(N-1)$,

(iii) $D_H = [R - E(R)]/S(R)$,

(iv) $D_A = [R - E(R)]/(\max R - \min R)$.

R, as before, is the observed number of repetitions in the series; $E(R)$ and $S(R)$ are respectively the value of R and the standard deviation of R that would be expected if the series were randomly arranged; max R and min R are respectively the maximum and minimum possible values of R for that particular series. Formulae for computing these terms are given in Hudson & Dunn (1969) and Dalrymple-Alford (1970). In the absence of established usage, I have followed Dunn (1969) in using D for these difference measures, with distinguishing subscripts. D_B is simply the difference between the observed and expected number of repetitions, a measure that appears to be recommended by Bousfield & Bousfield (1966). D_R is the difference between the observed and expected values of RR index, and appears to be

one of W. A. Bousfield's early suggestions for measuring clustering. D_H is the Hudson–Dunn measure, while D_C is the difference between observed and expected C' values, used first by Dalrymple-Alford & Aamiry (1969).

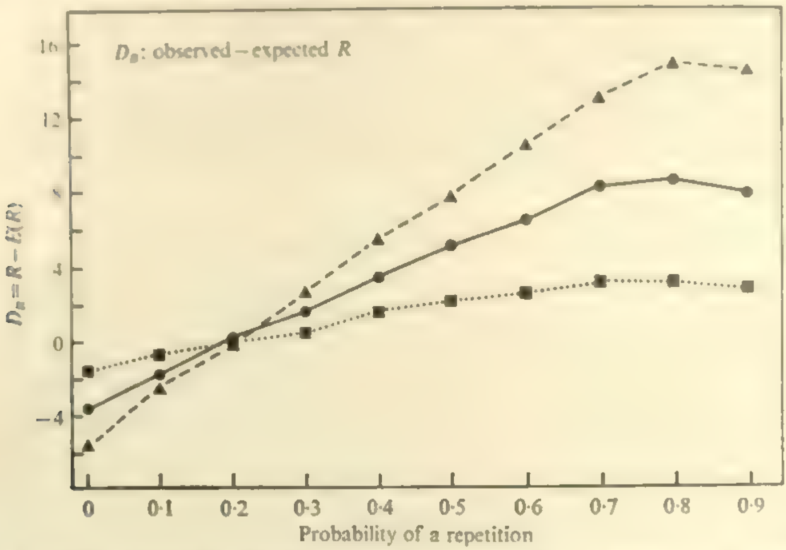


Fig. 2. D_R as a function of sequence length and probability of a repetition, for computer generated sequences. Sequence length: ■ . . . . . ■, 10; ●——●, 20; ▲- - -▲, 30.

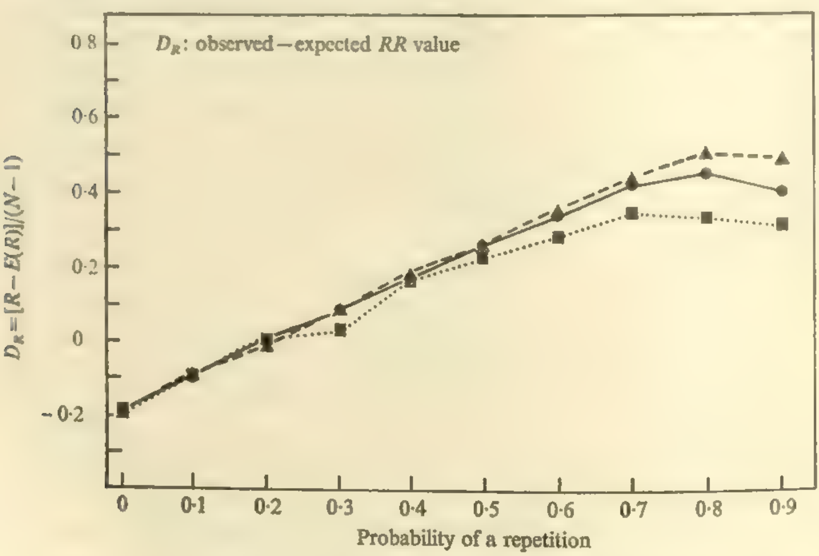


Fig. 3. D_R as a function of sequence length and probability of a repetition, for computer-generated sequences. Key as in Fig. 2.

Mean values for each of these measures were computed for each length of series in each of the ten sets, after eliminating sequences consisting of only one type of letter, i.e. cases of indeterminate clustering. Interest lay in the extent to which the

values of each measure varied with p, but remained constant over sequence length for a fixed value of p. Figs. 2–5 summarize the findings. It appears that neither D_R nor D_H are satisfactory measures for they are sensitive to the length of the sequence and not only to the tendency for items to cluster. D_I is better, but for moderate levels of clustering — at high levels it appears also to be sensitive to sequence length. D_A fares best of all, though for moderate levels of clustering there is a tendency for higher values to occur for shorter sequences.

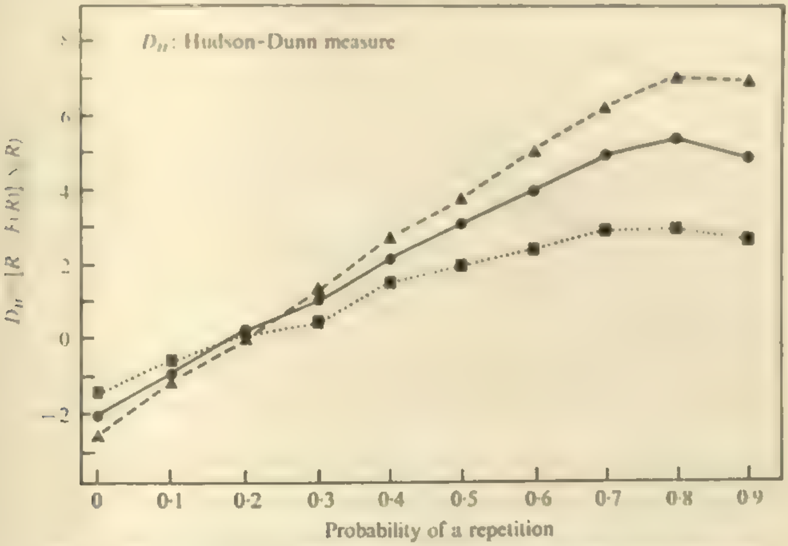


Fig. 4. D_H as a function of sequence length and probability of a repetition, for computer-generated sequences. Key as in Fig. 2.

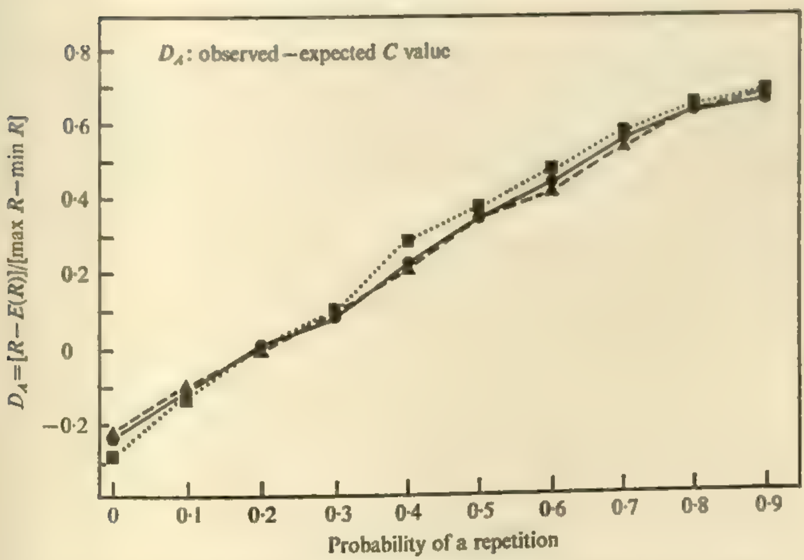


Fig. 5. D_A as a function of sequence length and probability of a repetition, for computer-generated sequences. Key as in Fig. 2.

Until a way can be found to generate sequences consisting of different numbers of categories but the same degree of clustering, the evaluation of these measures must be considered incomplete. However, the implications of the present exercise are fairly clear. Of the measures examined here,* D_A is least likely to mislead, while it seems that conclusions based on D_B, D_E and D_H must be regarded with suspicion if the data show that higher values are associated with greater output, for this may be nothing more than the bias of these measures. We are left then with the disconcerting thought that many of the results reported so far might after all be artifactual.

I am indebted to Dicran Keosheyan of the University Computer Center for writing the computer program.

REFERENCES

BARTON, D. E. & DAVID, F. N. (1957). Multiple runs. *Biometrika* **44**, 168–178.

BOUSFIELD, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *J. gen. Psychol.* **49**, 229–240.

BOUSFIELD, A. K. & BOUSFIELD, W. A. (1966). Measurement of clustering and of sequential constancies in repeated free recall. *Psychol. Rep.* **19**, 935–942.

BOWER, G. H., LESGOLD, A. M. & TIEMAN, D. (1969). Grouping operations in free recall. *J. verb. Learn. verb. Behav.* **8**, 481–493.

COFER, C. N. (1967). Does conceptual organization influence the amount retained in immediate free recall? In B. Kleinmuntz (ed.), *Concepts and the Structure of Memory*. New York: Wiley.

DALRYMPLE-ALFORD, E. C. (1970). Measurement of clustering in free recall. *Psychol. Bull.* **74**, 32–34.

DALRYMPLE-ALFORD, E. C. & AAMIRY, A. (1969). Language and category clustering in bilingual free recall. *J. verb. Learn. verb. Behav.* **8**, 762–768.

DUNN, J. E. (1969). A compounded multiple runs distribution. *J. Am. statist. Ass.* **64**, 1415–1423.

HUDSON, R. L. & DUNN, J. E. (1969). A major modification of the Bousfield (1966) measure of category clustering. *Behav. Res. Methods & Instr.* **1**, 110–111.

SHUELL, T. J. (1969). Clustering and organization in free recall. *Psychol. Bull.* **72**, 353–374.

* Since the MRR index is identical to the C index when min R is zero, and since min $R > 0$ is relatively infrequently encountered in practice, the D_A measure has a mean value that differs slightly from what one would obtain with a difference measure defined as MRR $- E$(MRR), i.e. $[R - E(R)]/\max R$. For this reason the latter measure was excluded from this exercise. However, subsequent analysis of the same sequences was carried out and showed that there were no differences between D_A and this measure up to $p = 0.4$. From then on the curves for this measure flatten out to a slightly greater extent than do the curves for D_A. There was little to choose between the two measures as far as variation with sequence length for a fixed value of p was concerned.

(*Manuscript received* 19 *May* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 335–346
Printed in Great Britain

RECOGNITION MEMORY FOR TARGETS FROM A SCANNED WORD LIST

By ARTHUR I. SCHULMAN
Department of Psychology, University of Virginia, Charlottesville, U.S.A.

A word list was designed so that half its words would denote targets when any of a number of target classes were defined. After scanning this list for targets, subjects were unexpectedly tested on their ability to recognize the words they had scanned. Recognition memory depended on how targets had been defined: a search for members of a semantically defined target class resulted in better recognition than a search for targets structurally defined. As a group, targets were better recognized than non-targets after a search for words that (a) contained the letter A, (b) denoted living things, or (c) denoted geographical locations. Recognition memory declined by as much as [illegible] per cent from the first quarter of testing to the last. Items prior to a word's input as well as decisions prior to its testing diminished its recognizability, whereas items following its input exerted comparatively little deleterious effect.

Suppose that a long word list is scanned for targets within it, and that all words then are re-presented unexpectedly in a test of recognition memory. If the ability to recognize items recently scanned depends on what can be remembered about the encounter, then certain types of search should lead to better recognition than others.

A search for structurally defined targets should produce a poorer basis for later recognition than a search for targets semantically defined. Two examples will illustrate this point. Deciding that a word is a member of the target class of living things must involve not only the identification of the word as a word, but also a check of some of its stored semantic attributes. But for targets defined as words that contain the letter A, deciding that a word is a target need not entail any semantic processing whatsoever.

The rationale behind the present research, then, is that how a target is defined determines the way in which a scan will be carried out; that what one does when looking for targets provides the basis for later recognition; and that a better basis is established by a search that involves semantic processing than by one that examines only structural features of the input.

The search for targets, moreover, may result in different treatment of targets and non-targets, so that recognition memory for targets may prove to be superior to that of non-targets. Indeed, it has been suggested (Neisser & Beller, 1965) that, when a subject scans a list in search for a randomly placed target word, he makes no lasting record of the words through which he searches. Neisser & Beller argue that, even when a target is defined by membership in a semantically determined class, words implicitly rejected as targets are 'not stored, nor even registered', and are therefore not recognized when tested unexpectedly. This plausible argument is somewhat weakened, however, by the fact that their subjects failed to discover the target they sought on 17 per cent of all lists scanned. Such a high percentage of misses may reflect gaps in the scan that inevitably would result in the non-recognition of unscanned words.

The present research, unlike Neisser & Beller's, employs a word list with large numbers of targets and non-targets and requires a decision by the subject about each word. With the subject's attention directed to each word, it would be surprising if our non-targets left no trace useful for later recognition. Nevertheless, the requirements of the scanning task might under certain target conditions lead to better recognition of targets than of non-targets.

The experiments described below were therefore designed with two main objectives in mind: to compare recognition performance after prior search for targets either structurally or semantically defined and, however a target was defined, to compare recognition of target and non-target words.

METHOD

The scanning list. A master set of 192 familiar English words was prepared. All but 15 occurred at least once per million words in the Thorndike-Lorge (1944) G count, and only four did not appear at all in that compilation. All words were from five to eight letters long. Every word either contained the letter A or did not; either had a repeated letter or did not; and denoted either a living thing or a geographical location. Half the words independently contained an A, had no repeated letter, and denoted a living thing. Each word thereby fell into one of $2^3 = 8$ mutually exclusive and exhaustive subclasses, with 24 representatives of each subclass. Half the words in each subclass were chosen randomly for inclusion in the scanning list. Those that remained were earmarked as 'new' words for the recognition test that was to follow the scan. The scanning list thus consisted of 96 words, 12 from each subclass. The list was constructed so that each block of eight words would contain exactly one member of each subclass, with the proviso that no four consecutive words be drawn from the same one of the six large classes (e.g. living things, words with no A, etc.). The first block of words scanned by all subjects – QUAIL, BRAZIL, TRENTON, PARROT, TUNIS, TIGER, WARSAW, TERMITE – exemplifies the structure of the list.

The scanning task. The scanning list was presented in booklet form, 48 words to each of two pages, and 16 words to each of three columns per page. The words were typed, using capital letters only. For all groups but one, the letters Y and N appeared to the right of each word. Subjects knew that half the words to be scanned would be targets. They were instructed to circle the Y for targets, the N for non-targets, and to work as fast as possible consistent with accuracy. Total scanning time was recorded (to the nearest 5 sec.) by the subject himself: upon concluding his scan he simply copied the elapsed time that the experimenter had just written on the blackboard (all subjects having begun the scan simultaneously). Thus it was possible to determine the average time spent on a word under each target condition.

Five groups of subjects were differentiated only by the target instructions they received. Group A looked for words that contained the letter A; group D for words all of whose letters were different; group L for words denoting living things; group G for words denoting geographical locations; and group LG, both for living things and for geographical locations. For group LG all words were therefore targets, and the subject had to choose the semantic target class into which each word fell. He recorded his decision by encircling L or G next to each word rather than Y or N.

The recognition test. No group was told that a recognition test would follow the scanning phase of the experiment. All were led to believe that the experimenter's interest was in the speed and accuracy of their scanning. When all subjects in a group had completed the scanning phase, they were told matter-of-factly that they would be tested on their ability to recognize the words on the list they had just scanned. After brief instructions, a forced-choice recognition test was administered, each choice being between a word that was on the scanned list and one that was not. The 'new' and 'old' words of a pair were both of the same subclass, e.g. a scanned word that contained an A, that had at least one repeated letter, and that denoted a living thing would be paired with a new word with the same set of properties. The test consisted of 96 such pairs so that all scanned words, both targets and non-targets, would have an opportunity to be recognized. Subjects worked at their own pace, but under instructions not to spend more than a second or two on each decision. For half the 96 test pairs, the scanned word was on the left. The test

[illegible] was arranged so that words that appeared early (late) in the scanned list also were tested early (late) during the recognition phase. There was a [illegible] test pairs [illegible] test position. Each block of eight test pairs included exactly one pair from each of the eight [illegible] described earlier. The first block of the recognition test exemplifies these properties: TERMITE, SQUIRREL, WARSAW, TRINIDAD, DOLPHIN, STURGEON, RHINE, EGYPT, [illegible] GAZELLE, NORFOLK, OREGON, MAINE, AZORES, LIZARD, MAGPIE. The same recognition test was administered to all experimental groups.

Subjects. All subjects within a group were run in the same session. There were 16 subjects in each of groups A, D and L, 12 subjects in group LG, and 10 subjects in group G. Students from an introductory course in psychology served as subjects, thereby fulfilling one of the course requirements.

Results and discussion

Performance on the recognition test will be indicated by the proportion of correct decisions, $P(C)$. Table 1 presents the main results. A one-way analysis of variance—using the *numbers* of correct decisions, not the tabulated proportions, as data—confirms the apparent significance of the obtained differences among the experimental groups ($F = 17{\cdot}3$; d.f. $= 4, 65$; $P < 0{\cdot}001$). Orthogonal comparisons of interest put group A against group D ($t = 3{\cdot}30$; d.f. $= 30$; $P < 0{\cdot}01$), groups A and D against

Table 1. *Overall recognition performance, target v. non-target recognition,* and scanning times*

Target condition ...	A ($n = 16$)	D ($n = 16$)	L ($n = 16$)	G ($n = 10$)	LG ($n = 12$)
Overall $P(C)$	0·677	0·754	0·811	0·835	0·859
Standard deviation	0·077	0·071	0·057	0·048	0·051
$P(C)$, targets	0·730	0·766	0·863	0·877	0·884 (L)
Standard deviation	0·107	0·096	0·062	0·051	0·055
$P(C)$, non-targets	0·626	0·742	0·759	0·794	0·833 (G)
Standard deviation	0·110	0·086	0·078	0·076	0·064
Scanning time per word (sec.; includes time for motor response)	1·00	2·28	1·16	1·08	1·15
Standard deviation	0·15	0·33	0·22	0·16	0·22

* Each value of overall $P(C)$ is based on $96n$ decisions. Each value of $P(C)$ for targets, as well as each value of $P(C)$ for non-targets, is based on $48n$ decisions. Although values of $P(C)$ and their standard deviations are tabulated here, statistical analyses reported in the text make use of obtained frequencies, not proportions, of correct judgements.

groups L, G and LG ($t = 7{\cdot}32$; d.f. $= 68$; $P < 0{\cdot}001$); group L against group G (no difference anticipated, $t = 0{\cdot}90$; d.f. $= 24$; $P > 0{\cdot}30$); and groups L and G against group LG ($t = 1{\cdot}61$; d.f. $= 36$; $P > 0{\cdot}10$). The most obvious result is that groups A and D, for whom targets were defined structurally, differ reliably from each other as well as from the groups for whom targets were semantically defined. That group LG fails, at conventional significance levels, to outperform groups L and G collectively is somewhat surprising. It should be noted, however, that for half the words scanned, the task requirements for groups LG and L and for groups LG and G were presumably the same. Only those words that are processed differently during the scan might be expected to differ in their recognizability, and we shall see below that this appears to have been the case.

Target v. non-target recognition

For all groups except LG (for whom all words were targets), the scanning list consisted of 48 targets and 48 non-targets. Because the test of recognition paired each old word with a new one drawn from the same subclass, recognition memory for targets and for non-targets may be compared. Table 1 shows that there were large differences in favour of targets under conditions A (t = 4·44; d.f. = 15; P < 0·001), L (t = 4·81; d.f. = 15; P < 0·001) and G (t = 2·84; d.f. = 9; P < 0·02), but no difference under condition D (t = 0·95; d.f. = 15; P > 0·30). Under condition LG, the difference in recognition memory between living things and geographical locations, though unexpected, is also statistically significant (t = 2·72; d.f. = 11; P < 0·05). This difference may not indicate that animals are easier to recognize than places – compare the results of conditions L and G – but only that some subjects proceeded as if under group L instructions during the scan. If a word were not a living thing, the G-response could then be chosen without the word being positively identified as a geographical location.

Semantically defined targets

The data for conditions L and G are particularly interesting. When a subject scanned for living things he recognized them more easily than geographical locations; but when he scanned for geographical locations, the reverse was true. In addition, group LG recognized geographical locations more easily than group L (t = 2·58; d.f. = 26; P < 0·02) and recognized living things more easily than group G (t = 3·04; d.f. = 20; P < 0·01). For at least those groups with semantically defined target classes, then, positive identification of a target appears to provide a better basis for later recognition than a decision that a word is *not* a member of a target class.

Structurally defined targets

Mere 'targetness', however, proved to be of no special benefit under condition D. In looking for words without a repeated letter, it is hard to see how, except for the classification response, targets and non-targets might be processed differently. Without differential processing of targets and non-targets there is no reason to anticipate differences in their later recognizability. Going a step further, one might argue that a structural search would favour target recognition only when the processing it entailed fostered the encoding of targets – but not of non-targets – as perceptual units. Mandler *et al.* (1969) implicitly acknowledge this argument in the word-sorting task they required of their subjects, who were not permitted to sort on the basis of structural features.

Condition A produced far better recognition of targets than of non-targets, and it is easy to argue that this condition, unlike condition D, made it possible to process targets and non-targets differently. The search for an A may involve pre-perceptual processes (Neisser, 1967), so that a word may not even be perceived as such unless evidence is found for an A within it. Since such evidence is more likely to be yielded by words than indeed contain A's than by words that do not, targets under this condition should be better recognized – as they were – than non-targets. If this analysis is correct, it may be easier to understand why positive classification alone

might enhance recognition for semantically defined targets, but not for structurally defined targets. For semantic targets, word perception must precede the identification of a word as a target; the identification then serves as an attribute of the word. For structural targets, target identification may precede perception of the word and not serve the attributive function (cf. Wickens, 1970; Underwood, 1969) that is important for recognition.

Under condition A, the recognition of targets depended on the location of the letter A within the target word. Fig. 1 shows that the earlier the A, the more likely the recognition of the word in which it appeared. If A-detection indeed precedes word-perception, this decreasing function suggests that a subject scans a word from left to right, and tends to perceive only that portion of a target to the right of the location of its A.

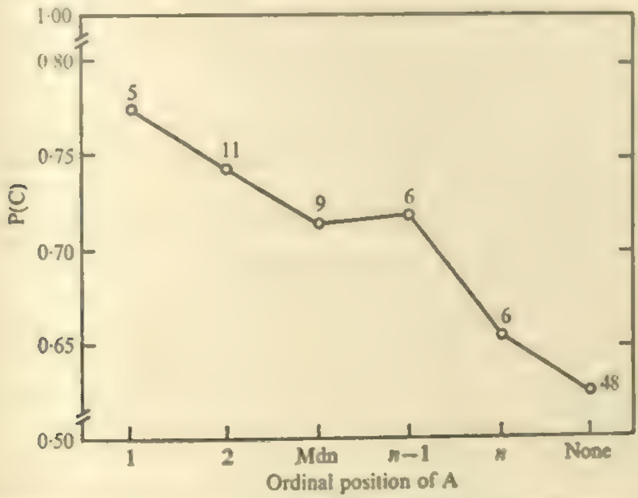


Fig. 1. Recognition memory for targets under condition A (n = 16) as a function of location of the letter A within the target word. Words with more than one A were ignored. The figure near each point represents the number of words, x, whose A appeared in the indicated ordinal position. Thus the proportion of correct decisions, $P(C)$, is based on $16x$ decisions. On the abscissa, 'Mdn' denotes a median or near-median letter position within a target word; near-median letter positions include the third or fourth letter of a six-letter word. 'None' denotes the entire set of non-target words.

Structural targets v. semantic non-targets. Table 1 reveals that overall $P(C)$ under condition D, $P(C)$ for targets under condition A, and $P(C)$ for non-targets under conditions L and G were similar in value. The possibility exists that these comparable levels of performance derived from the same degree of semantic processing for the words in question. This would mean that targets under condition A and all words under condition D were likely to be perceived as words, but were unlikely to receive additional semantic processing; and that non-targets for groups L and G, after being perceived as words and failing the target test, were equally unlikely to receive additional semantic processing. These possibilities deserve further exploration.

Scanning rates

Table 1 also records the mean scanning time per word for all main experimental groups. These times include the 0·45 sec. or so required to make the motor response of circling the appropriate letter. Within each group there was no correlation between a subject's scanning time and his $P(C)$. The scanning task imposed by condition D consumed by far the most time, whereas scanning for words that contained the letter A took only slightly, but significantly ($P < 0·05$, Mann–Whitney test), less time than scanning for semantically defined targets. Condition LG, which required that all living things and geographical locations be identified as such, was no more time-consuming than either of conditions L or G, which required only yes-or-no decisions. This suggests that separate checks on whether a word is an animal or a place may be carried out by parallel processes. Although their scan took no longer, group LG, as already noted, outperformed group L in recognizing geographical locations, and outperformed group G in recognizing living things. Clearly condition LG fostered more effective use of scanning time than either of conditions L and G. All in all, the scanning times indicate that what is important for recognition is not simply the duration of the scan, but what the subject does while scanning.

Errors made while scanning and their consequences for recognition memory

Very few classification errors were made during the scan. With false positives and false negatives pooled, the proportion of errors was 0·010 under condition A, 0·033 under D, 0·027 under L, and 0·027 under G. Under condition LG, the proportion of mistaken classifications was 0·010. Despite the paucity of scanning errors, they merit analysis because of insights they might afford into recognitive processes. The ability to recognize a word mistakenly classified should depend more on the reasons for the misclassification than on the mere fact of its occurrence. If the steps leading to a word's classification were known, the chances of its later recognition might better be forecast.

A close look at one type of classification error – false negatives under condition L – proved instructive. Thirty of the 41 scanning errors made by the 16 subjects in group L were false negatives. Of these 30, seven were idiosyncratic. The remaining 23 represented six words – mastiff, anchovy, gibbon, ermine, sturgeon and iguana – mistakenly classified by at least two subjects. If the steps that led to false negatives were the same as those that led to rejection of true non-targets, these six words should have been as poorly recognized as all non-targets. But what if such mistaken classification were symptomatic of special treatment these words received during the scan, owing to their own peculiar properties and the type of search imposed by condition L? Under these circumstances the six should have proved easy to recognize not only for those subjects who overtly misclassified them, but for group L as a whole. This was indeed the case: of the 96 responses (6 words × 16 subjects), 92 were correct, yielding $P(C) = 0·96$. Such extraordinary ability to recognize targets falsely classified should be compared with overall target recognition for this group, whose $P(C) = 0·863$. The ease of these words' recognition is all the more remarkable in view of the fact that most of them were tested for recognition during the second half of the test when – as will be seen in the next section – overall $P(C)$ was much lower than in the first half of testing.

It is not simply that these words are intrinsically easy to recognize. They were indeed easy, as one might expect, for the subjects of group LG: $P(C) = 0{\cdot}94$ c. $P(C) = 0{\cdot}884$, overall, for living things. But for all other groups recognition of these same six words was no better than average.

It is hard to escape the conclusion that these six words received special treatment by group L during the scan. For some, e.g. 'anchovy' and 'ermine', the special treatment could derive from their 'competing' membership in semantically defined classes other than that of living things. For others, e.g. 'mastiff' and 'iguana', there may have been uncertainty owing to unfamiliarity, as to whether the word belonged to the target class at all. In all of these cases, it is suggested, recognition is facilitated when one can recall one's uncertainty in classification during the scan. That recognition of the six words was only average under condition G suggests that, whereas condition L generated uncertainty as to whether these words were targets, no comparable uncertainty was generated under condition G in deciding that they were not.

The superior recognition that attends at least one type of classification error complicates the interpretation of the differences between targets and non-targets obtained under conditions L and G. Although large differences remain when the six false negatives are excluded from analysis, we cannot be sure that these differences are due to the supposed advantages that positive classification holds for recognition memory. In semantic searches, at least, a target's advantage may lie not merely in its positive tagging, but rather in the more distinctive processing it would tend to receive en route to being classified as such.

Other types of scanning error might also shed light on recognitive processes, but the remaining error groupings were too small to permit justifiable inferences. One obvious prediction, however, would be that false negatives under condition A should behave like true negatives, i.e. a word whose A goes undiscovered should be just as difficult to recognize as words without A's. An experiment that promoted scanning speed at the cost of reduced accuracy might yield enough false negatives to permit a test of this hypothesis. Unlike the previous example, however, a classification error during this sort of structural search would probably be idiosyncratic, and would have no implications for group performance.

Finally, the 50 scanning errors under condition D included only six where a subject reported a repeated letter when none was present, and 44 where a subject failed to notice that a word contained a repeated letter. The latter's frequency of occurrence increased with the actual separation between those letters that were repeated within the word. For example, 37 of these 44 errors were made with the 28 words whose repeated letters were separated by at least two others, while only seven errors of this kind were made with the 20 words whose repeated letters were separated by no more than one letter. There was no evidence, however, that scanning errors of either type affected recognition memory under this experimental condition.

The declining course of recognition memory

The values of $P(C)$ reported thus far are measures of average performance on the 96-pair recognition test. They conceal the fact that ability to recognize the scanned words fell off sharply as testing proceeded. Fig. 2 presents the course of recognition memory over 24-pair blocks of our test sequence, and shows that d' – the

discriminability parameter derivable from $P(C)$, as explained in Swets (1964)) was reduced by an astonishing 55–75 per cent from the first quarter of testing to the last.

Remember that testing was conducted on a first-in, first-out basis, with a correlation of 0·93 between input and test position. There was therefore a nearly constant lag between a word's input and its test position. The uniformity of the lag, coupled with the high initial levels of performance, suggests at most a small inhibitory effect of late-scanned words on the recognition of words that preceded them (more on this later). The decline in d' over the course of the recognition test would seem, then, to depend largely on the combined effects of two thoroughly confounded variables: the number of items prior to a word's input and the number of decisions prior to its testing.

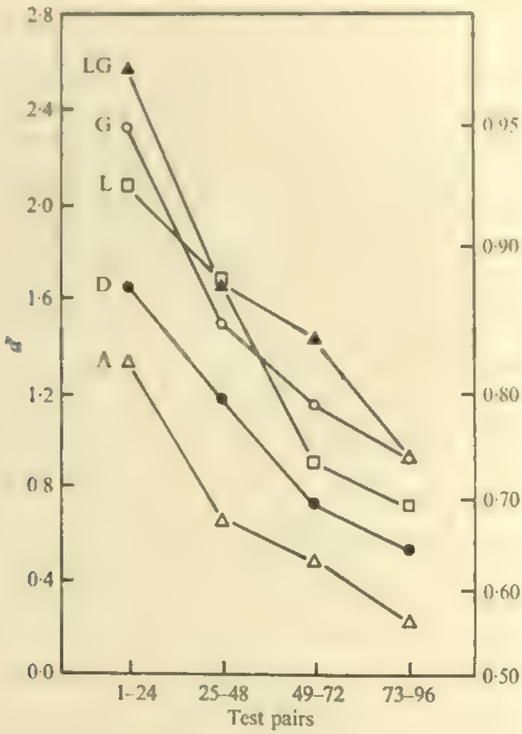


Fig. 2. The declining course of recognition memory.

Proactive interference with recognition

By reversing the test sequence, the relative influence of these two kinds of pro-action may be assessed. If the decline in d' over the test were due solely to an increase in the number of prior decisions, testing for recognition on a first-in, last-out basis should produce the same 'forgetting functions' as those of Fig. 2. But if the number of prior inputs were all that mattered, a reversed test sequence would produce functions the mirror-image of these, with $P(C)$ increasing over the test. A group of 15 subjects was tested under LG instructions, but with the original test sequence reversed. (I thank M. A. Browne for her assistance here.) Overall $P(C)$ for this group was 0·867, nearly the same as the $P(C) = 0·859$ of the original LG group. Their

forgetting curve, however, was nearly flat ($P(C)$ = 0·87, 0·87, 0·85 and 0·88 for the four quarters of testing), indicating that prior inputs and prior decisions exerted nearly equal deleterious effects on recognition memory. With a perfect negative correlation between input and test position, a flat function would signify that a constant sum of prior inputs and prior decisions always produced the same value of d'. Procedural changes (slowing the scanning rate or not requiring target identification, to mention but two) might of course be expected to alter the relative effects of input and output proaction.

The proactive effect during input may result from later words being processed less distinctively than earlier ones, owing to the progressive 'automatization' of the task. As the subject progresses through the scanned list, he may grow more skilled in classification, so that late words are processed more automatically, and perhaps faster, than early ones. Increased efficiency is ordinarily accompanied by decreased awareness of what one is doing, but it is precisely the awareness of the steps involved in classification that may be essential for recognition. This may help to explain the near-perfect recognition of false negatives under condition L described in the previous section. A word may be recognized as having appeared in a recent context only if one can remember the encounter. The more automatically a word is processed during the scan, the poorer the cues available for its later recognition.

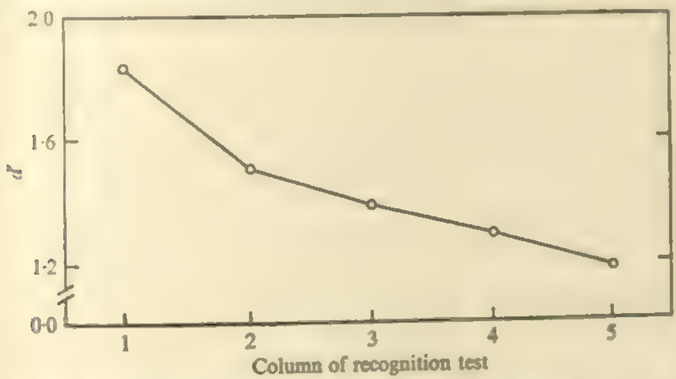


Fig. 3. The declining course of recognition memory for nonsense syllables. Derived from Peixotto (1947); see text for explanation.

The proactive effect of earlier decisions upon the recognition of subsequently tested items is more difficult to account for, and no attempt to do so will be made here. But however such proactive effects are produced, the experimental literature on recognition memory does not adequately prepare us for them. Only one earlier study (Peixotto, 1947) reports within-test proaction that may be compared with ours. Her subjects looked at 90 nonsense syllables, 30 of which were later to appear among 90 new syllables in a test of recognition. A yes–no recognition test was used, with 24 syllables in each of five columns. For each column, Peixotto reports the proportions both of old and of new syllables judged to be old. Each such pair of proportions may be translated into a single value of d': the forgetting function that results is presented in Fig. 3. Although not even half so steep as the curves of Fig. 2 – Peixotto allotted 2 sec. per word during input and seems to have used a haphazard test sequence in

which an 'old' word's test position was uncorrelated with its position at input – the proactive effect is clear.

In apparent contrast to these forgetting curves are the findings of Norman & Waugh (1968). In their Expt. I, subjects were presented with 30-word input lists that alternated with 20-word yes–no tests for 45 cycles. (The input rate was 1 sec. per word, the test rate 2 sec. per decision.) Recognition memory for a word appearing early in the immediately preceding input list was found to be independent of its test position. In other words, first-in, first-out testing produced no better recognition than first-in, last-out testing – a result at variance with the present findings.

Shepard & Chang (1963) and Shepard & Teghtsoonian (1961) used procedures that mingled inputs and outputs – Norman & Waugh (1968) alternated them and obtained 'steady-state' recognition that, in the former case, had stabilized after 16 decisions or so. In all of these experiments, both the number of prior inputs and prior decisions increased as the subject progressed through the test session. Their cumulative effect might be expected eventually to produce asymptotically stable performance, as even the decelerating functions of Fig. 2 suggest. Given the nature and number of the stimuli these studies employed, as well as the rate and type of processing their tasks entailed, their performance measures may reflect predominantly asymptotic recognition that would overwhelm early proactive effects. Clearly, further research is required to define those conditions that engender proaction and those that do not.

The unimportant role of retroactive interference

While an increasing number of prior inputs and prior decisions both served to reduce the probability that a word would be recognized, the number of words subsequent to its input had apparently little effect on its recognizability. Under condition LG, for example, overall $P(C) = 0{\cdot}965$ for the first 24 test-pairs. Half the subjects in this group made no recognition errors at all during this first quarter of testing. Such error-free performance suggests that, once a word is classified during the scan, its long-term recognition is affected only slightly by the processing and storage of further input.

The absence of a substantial retroactive effect of input upon recognition was first reported by Heine (1914). But the literature has been ambiguous on this point, and the conditions under which retroaction is unimportant remain poorly understood. Shepard (1967), for example, reports somewhat better recognition for words from the second half of an inspection sequence than from the first half: apparent evidence for even greater retroactive than proactive effects upon memory. Shepard's inspection list, however, was extremely long – 540 words – while his test consisted of only 60 pairs. Since words from the inspection sequence appear to have been selected at random for testing at apparently random points of the test sequence, Shepard's results may have been determined largely by the asymptotic recognition of words from late in the inspection sequence.

In order to further explore proactive and retroactive phenomena, two additional groups of subjects ($n = 11$ in each group) were run. (M. A. Browne ran the subjects.) Under LG instructions they scanned an expanded list of words before being administered the same recognition test used in the main experiment. Group 96–24

scanned the original 96 words followed by 24 others drawn from the same population, while group 24–96 scanned the original 96 words after these additional 24 words had first been scanned. The 24 words were dummy items, and were not tested for recognition.

Two of the findings deserve mention. First, the performance of the 24–96 group ($P(C) = 0{\cdot}806$), but not that of the 96–24 group ($P(C) = 0{\cdot}842$), was significantly poorer than for group LG ($t = 2{\cdot}31$; d.f. $= 21$; $P < 0{\cdot}05$). Secondly, the relatively poor performance of group 24–96 was due almost entirely to degraded recognition on the first half of the recognition test ($P(C) = 0{\cdot}872$ compared to $P(C) = 0{\cdot}913$ for group 96–24; $t = 2{\cdot}17$; d.f. $= 20$; $P < 0{\cdot}05$). On the second half of the recognition test, differences between groups 96–24 and 24–96 ($P(C) = 0{\cdot}771$ and $0{\cdot}740$, respectively) are not statistically reliable ($t = 0{\cdot}95$; d.f. $= 20$; $P > 0{\cdot}20$). These results support the contention that a word's recognition depends more on the number of words preceding it than following it in the scan. The processing of words late in the scan appears to have little effect either on the retention of earlier words or on the distinctiveness of their cues; the processing of early items, however, may result in relatively ineffective cues being associated with subsequently scanned words.

Summary and conclusions

1. Recognition memory for targets from a previously scanned word list depends critically on how those targets were defined. When a subject searches for semantically defined targets, his later ability to recognize all words scanned is greater than after a search for targets structurally defined.

2. Targets may or may not be better recognized than non-targets, depending on how a target is defined. Only if a target's definition dictates that targets and non-targets be processed differently should we expect differences in recognition memory for targets and non-targets. Thus, words containing the letter A were better recognized than words that did not, presumably because the target words were more likely to be perceived as whole words once evidence for an A was discovered. But a search for words with no repeated letter resulted in no difference in the recognition of targets and non-targets. Animals were recognized more easily than places after a scan for animals, while the reverse was true after a scan for places. These semantically defined targets may have benefited merely from their positive identification as targets or else from the distinctive processing that targets may receive before being classified as such. The fact that targets can be better recognized than non-targets may call for modifications of the frequency theory of verbal discrimination, which at present (Underwood, 1970, personal communication) cannot account for such differences.

3. The time consumed by the scan is by itself a poor predictor of later recognition; how one uses that time is more important. The positive identification of a semantically defined target does not necessarily consume more time than a positive identification of a structurally defined target, but the former provides a much stronger basis for later recognition.

4. Few classification errors were made during the scan. The occurrence of one type of error, however, virtually guaranteed later recognition of the misclassified words. The nearly error-free recognition of false negatives under condition L suggests that

the processing demanded by this condition resulted in very special treatment of these words during the scan. Generally speaking, knowing the reason for classification errors during the scan should allow us to predict whether an error will affect the likelihood that the word mistakenly classified will later be correctly recognized.

5. The ability to recognize the scanned words declined by 55–75 per cent in d' from the first quarter of testing to the last. Similar 'forgetting curves' were obtained for all of the five experimental groups. First-quarter recognition was surprisingly good – $P(C) = 0{\cdot}965$ for group LG – under the first-in, first-out testing procedure that was used. Items prior to a word's input as well as decisions prior to its testing diminished its recognizability, but items following its input exerted comparatively little deleterious effect.

References

HEINE, R. (1914). Über Wiedererkennen und rückwirkende Hemmung. *Z. Psychol.* **68**, 161–236.

MANDLER, G., PEARLSTONE, Z. & KOOPMANS, H. S. (1969). Effects of organization and semantic similarity on recall and recognition. *J. verb. Learn. verb. Behav.* **8**, 410–423.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton–Century–Crofts.

NEISSER, U. & BELLER, H. K. (1965). Searching through word lists. *Br. J. Psychol.* **56**, 349–358.

NORMAN, D. A. & WAUGH, N. C. (1968). Stimulus and response interference in recognition-memory experiments. *J. exp. Psychol.* **78**, 551–559.

PEIXOTTO, H. E. (1947). Proactive inhibition in the recognition of nonsense syllables. *J. exp. Psychol.* **37**, 81–91.

SHEPARD, R. N. (1967). Recognition memory for words, sentences, and pictures. *J. verb. Learn. verb. Behav.* **6**, 156–163.

SHEPARD, R. N. & CHANG, J. (1963). Forced-choice tests of recognition memory under steady-state conditions. *J. verb. Learn. verb. Behav.* **2**, 93–101.

SHEPARD, R. N. & TEGHTSOONIAN, M. (1961). Retention of information under conditions approaching a steady state. *J. exp. Psychol.* **62**, 302–309.

SWETS, J. A. (ed.) (1964). *Signal Detection and Recognition by Human Observers*. New York: Wiley.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Bureau of Publications, Teachers College, Columbia University.

UNDERWOOD, B. J. (1969). Attributes of memory. *Psychol. Rev.* **76**, 559–573.

WICKENS, D. D. (1970). Encoding categories of words: an empirical approach to meaning. *Psychol. Rev.* **77**, 1–15.

(*Manuscript received* 8 *September* 1969; *revised manuscript received* 22 *July* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 347–352
Printed in Great Britain

THE INFLUENCE OF FORCE MAGNITUDE ON THE PERCEPTION OF BODY POSITION

II. EFFECT OF BODY POSTURE

BY H. SCHÖNE AND N. J. WADE*

Max-Planck-Institut für Verhaltensphysiologie, Seewiesen, Germany

With rotation in the sagittal plane, deviations of the perceived from the trained body position were measured at force magnitudes of 1·0, 1·4 and 1·8 g. In Expt. I body postures with corresponding utricular macular orientations in the region of 40° and 10° backward and 20° forward were examined. In the two backward positions the judgements were displaced forward with increased force operating, whereas a backward shift was produced for the forward utricular orientation. The results support the hypothesis that the position in which the body is perceived is related to the utricular shear force. In Expt. II body postures were investigated which involved orientations of the utricular maculae of about 70°, 90° and 110° backward. The small shifts in judgements produced at increased force magnitudes were not significant. It appears that the above hypothesis is only valid for shear values up to 1·0 g, and this may represent the physiological limit of the statolith system with respect to judgements of body position.

In the previous report (Wade & Schöne, 1971) the perception of body position with rotation in the sagittal plane was shown to vary with the force operating. In the first experiment the subject was trained to judge the upright body position accurately under normal gravity. At increased force magnitudes the position in which the subject felt upright was displaced forward relative to the trained position. The training position in the second experiment involved the same upright trunk posture but the head was inclined forward. Under these conditions the forward displacement of the judged body position was greatly reduced at increased force strengths. Since the inclination of the utricular statolith system (which is known to respond to linear accelerations) was reduced by forward head tilt, the displacement was attributed to an increase in the backward acting shear force on the utricular system. Thus, it was hypothesized that the perception of a particular body position corresponds to a specific shear force acting on the utricular maculae; modification of the shear force by means of centrifugation results in a change in the perceived body position.

The present experiments were designed to test the above hypothesis for body postures encompassing the horizontal and vertical orientations of the utricular maculae.

EXPERIMENT I

The upright position of the head is defined as that with the teeth firmly impressed in a horizontal bite-board. In this posture the utricular maculae are inclined backward (–) from the horizontal by approximately 40° (see Wade & Schöne, 1971). The three body positions investigated in this experiment were head upright (0°), and tilted forward (+) by 30° and 60°. The orientations of the utricular maculae were

* Present address: Department of Psychology, University of Dundee.

about $-40°$, $-10°$ and $+20°$, respectively. The first two head postures were examined in the earlier study, but in the present experiment the inclination of the head to the trunk remained unchanged for all head positions. According to the above hypothesis it was expected that the degree of forward shift in these two postures would be similar to those previously measured. However, with the bite-board tilted 60° forward the utricular maculae would be inclined forward. With increase in the force operating, the increment in the utricular shear would also be directed forward. Therefore, in order to equate the shear with that under normal gravity the body would need to be tilted backward relative to the trained position.

Method

The subject was trained to judge a particular body position accurately, and was then required to match his impression of this position during centrifugation. Apart from a different set of body positions, the only difference from the method employed in the previous study was a reduced number of force magnitudes. The same apparatus was used.

Subjects. Six subjects participated in the experiment.

Procedure. The three body positions examined corresponded to head positions with the bite-board horizontal, and tilted 30° and 60° forward. The trunk remained at the same inclination to the head for all postures; this relation was such that when the torso was upright the head was inclined 30° forward.

The training and test procedures were the same as those described in the first report (Wade & Schöne, 1971). The force magnitudes investigated were 1·0, 1·4 and 1·8 g. Each subject was tested in each posture. The orders for the six possible combinations of the three body positions were counterbalanced over subjects.

Results

The deviations of the perceived from the trained body position were averaged over sessions for each body position and force magnitude. The group means are represented in Fig. 1D for the three postures at each force magnitude. The individual data are shown separately for each posture. The overall trends for the curves differed significantly from one another ($F = 6{\cdot}32$; d.f. $= 4, 20$; $P < 0{\cdot}01$).

In the head-upright condition the body was shifted forward at increased force strengths in order to be perceived as the trained position. Such a shift was evident for all subjects but its magnitude differed widely between subjects (Fig. 1A). The forward displacement was reduced when the head was inclined 30° forward, and the individual differences were not as pronounced in this posture as with the head upright (Fig. 1B). These results are in agreement with those reported previously for the same head orientations (Wade & Schöne, 1971). Moreover, with the head tilted forward by 60°, so that the utricular maculae were also tilted forward, the judgements were displaced backward relative to those under normal gravity. This trend was clear for all but one of the subjects (Fig. 1C).

EXPERIMENT II

Body positions which involved orientations of the utricular maculae around the vertical were studied in this experiment. In such postures the utricular shear under normal gravity is around the normal limit of 1·0 g; centrifugation generates shear forces which exceed this value.

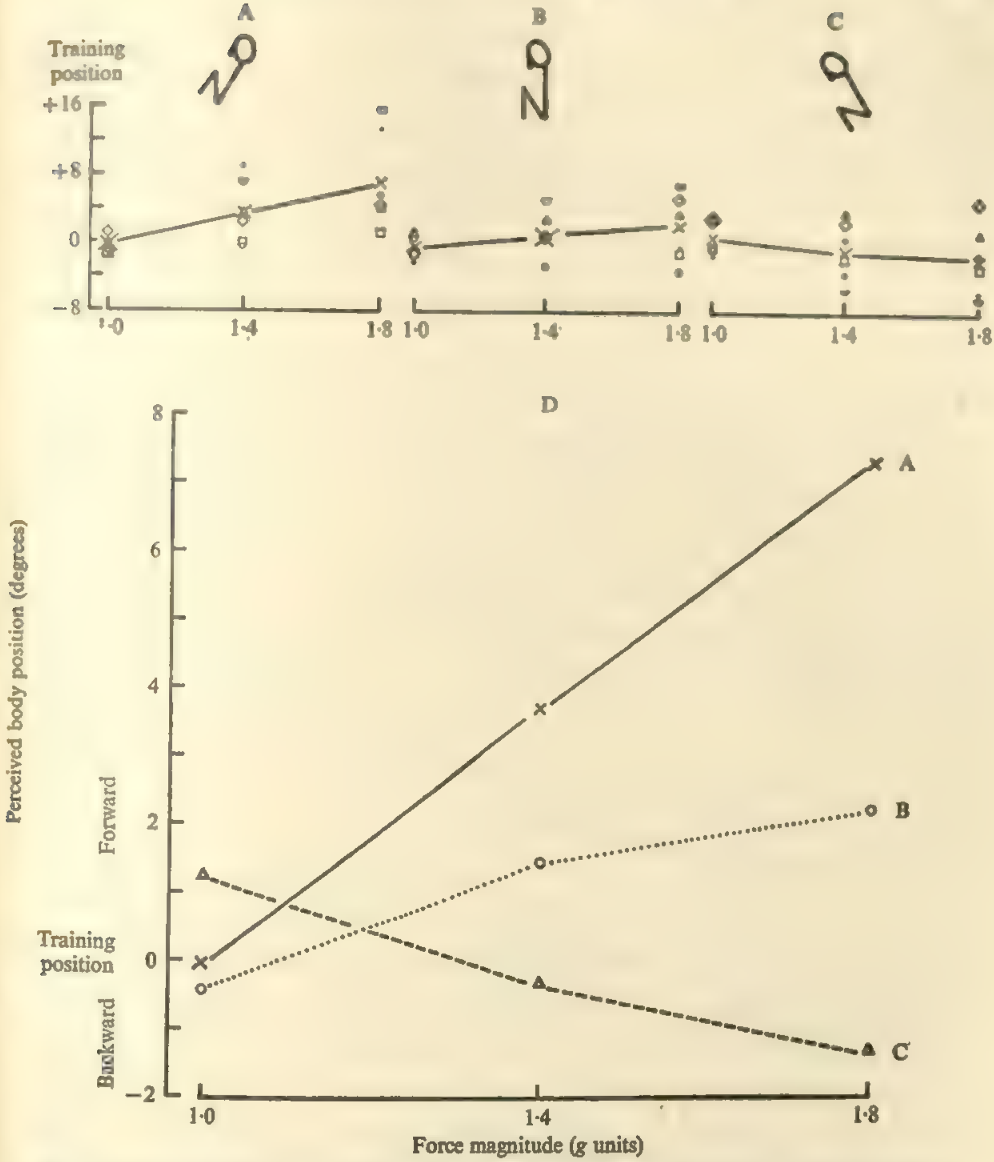


Fig. 1. Deviations of the perceived body position from the trained position as a function of the force magnitude for the different body postures. The upper graphs show the judgements for training positions with the head upright (A), tilted 30° forward (B), and 60° forward (C). The symbols represent the individual data in each posture. The group means are shown together in D.

Method

Due to apparatus constraints the range of starting positions was restricted to 10° and 20° forward and backward, and only four judgements were made per session instead of six.

Subjects. Six subjects participated in this experiment, three of whom had taken part in Expt. I.

Procedure. The three body positions investigated were backward head tilts of 30°, 50° and 70°. The corresponding backward inclinations of the utricular maculae were approximately 70°, 90° and 110°. The head-to-trunk relation was such that when the torso was upright the head was bent backward by 10°, and this remained constant for all head postures.

Results

The judgements of body position were averaged over sessions for each body posture and force magnitude. Fig. 2 shows the relationship between the force magnitude and the deviation of the perceived from the trained body position for the different body postures. The individual data for each posture are given separately. The overall trends for the curves did not differ significantly ($F = 1{\cdot}16$, d.f. $= 4, 20$, $P > 0{\cdot}05$).

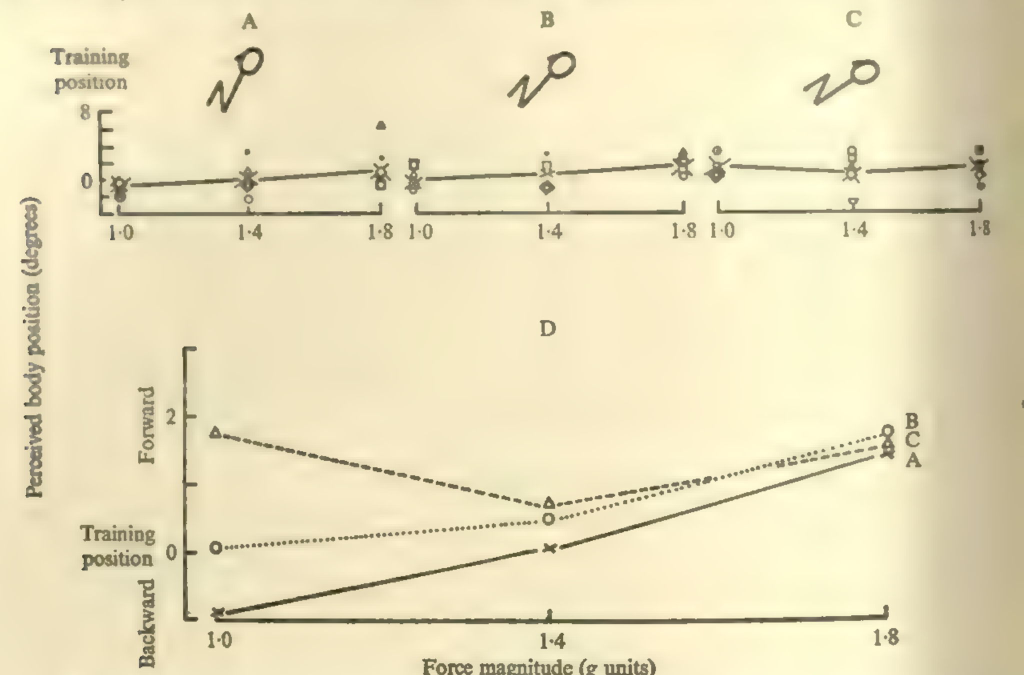


Fig. 2. Deviations of the perceived from the trained body position as a function of force magnitude for the different body postures. The upper graphs show the judgements for training positions with the head tilted backward by 30° (A), 50° (B) and 70° (C). The symbols represent the individual data in each posture. The group means are shown together in D.

With the head tilted backward by 30° for the trained position the judgements were displaced forward at greater force magnitudes. A smaller forward shift was produced for the condition with the head tilted 50° backward, and a slight backward shift was apparent for the 70° head tilt position. However, all these changes were small relative to the individual differences, and were non-significant statistically.

DISCUSSION

In both experiments judgements made under normal gravity showed systematic displacements from the trained position. In agreement with the results reported earlier (Wade & Schöne, 1971), judgements with the head tilted 30° forward were directed backward of the training position. This is the only posture that can be compared directly, as the trunk postures in the head upright conditions differed.

Nevertheless, in both instances with the head upright the judgements were backward with respect to the training position. With the head tilted 60° forward the judgements were forward of the trained position. Similar differences were evident in Expt. II. For the condition with the head tilted 50° backward the judgements were accurate but for head inclinations less than and greater than this degree the displacements from the trained position were backward and forward, respectively. The reasons for these deviations are not clear, as the training and testing conditions were the same. However, they may be associated with slight adaptation of the statolith system: the postures involving backward inclinations of the utricular maculae produced backward deviations, whereas those with forward (or beyond 90°) inclinations resulted in forward deviations.

Under increased force magnitudes in Expt. I the judgements of body position changed in the manner predicted. That is, with decreasing backward inclination of the utricular maculae the forward shift in perceived body position was decreased, and the displacement was reversed in direction for forward macular inclinations. The reliability of the effect is evident from the similarity of the curves with the head tilted 30° forward in both Expt. I and the earlier study (Wade & Schöne, 1971). A further point of interest concerns the comparison of the curves with the head upright in the two studies, as these were associated with different trunk orientations. The similarity of the values at different force strengths provides additional evidence that receptors in the trunk play a minor role in the perception of body position, at least under the conditions here investigated. However, while the utricular shear hypothesis received qualitative support from the above results, there were quantitative discrepancies – as was previously noted. That is, if the shear force operating in the training position was equated at increased force magnitudes then the values for the displacement in the perceived body position should have been greater than those obtained. For example, the predicted shifts for the three body positions from 1·0 *g* to 1·8 *g* are approximately +19°, +4° and −9°, whereas the obtained values were about +7°, +3° and −3°, respectively. It is possible that the discrepancies from the predicted values were due to the methodological factors detailed in the first report, as they were equally operative in the present experiment. Furthermore, the influence of other sensory systems, such as the saccular statoliths, cannot be excluded.

The methodological problems were probably intensified in Expt. II, because of the reduced range of starting positions employed. For instance, the likelihood of matching the time taken for rotation to and adjustment from the starting position was enhanced, and this would have reduced any shifts in the judgements. Although there was a slight change in the trend of perceived body position over force magnitude from forward to backward with increasing backward utricular tilt, this was not significant. The relatively large differences in judgements at 1·0 *g* were successively reduced by increases in the force operating. The results suggest that for judgements of body position the utricular shear hypothesis reaches its limiting condition for values of 1·0 *g*, which is the maximum normally produced with the maculae oriented vertically. The shear values with the maculae inclined by 70° and 110° were close to this limit, and the system appears to be relatively insensitive to further increases of shear in this region. A similar saturation effect has been demonstrated for visual verticality judgements with lateral body tilt during centrifugation (Schöne *et al.*,

1967). In addition, an inversion effect was found, i.e. with increased force magnitude the trend in the visual vertical for tilts greater than 90° was opposite to that with smaller degrees of body tilt (Schöne & Parker, 1967). Although there is a slight suggestion of such a shift in Fig. 2, there was no statistical verification for it. The detection of such small differences in these postures would require measurement over a wider range of conditions.

In conclusion, the results of Expt. I support the hypothesis that the utricular shear force influences the position in which the body is perceived. However, the degree to which the lack of direct correspondence between the value of shear and the perception of a particular body position represents the operation of additional factors or limitations of the method used remains to be determined. The marginal shifts in judgements that occurred for utricular orientations around the vertical (Expt. II) suggest that the hypothesis is only valid up to the normal extreme shear value of 1·0 g; this may be the physiological limit of the statolith system with respect to judgements of body position.

REFERENCES

SCHÖNE, H. & PARKER, D. E. (1967). Inversion of the effect of increased gravity on the subjective vertical. *Naturwissenschaften* **54**, 288–289.

SCHÖNE, H., PARKER, D. E. & MORTAG, H. G. (1967). Subjective vertical as a function of body position and gravity magnitude. *Naturwissenschaften* **54**, 288.

WADE, N. J. & SCHÖNE, H. (1971). The influence of force magnitude on the perception of body position. I. Effect of head posture. *Br. J. Psychol.* **62**, 157–163.

(*Manuscript received* 29 *June* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 353–365
Printed in Great Britain

THE DEVELOPMENT OF THE ABILITY TO DECENTER IN TIME

By RICHARD F. CROMER

Medical Research Council, Developmental Psychology Unit, Gordon Street, London

In order to study children's developing ability to decenter temporally, an experiment was designed analogous to Piaget's three mountain problem concerned with spatial decentering. Two types of temporal decentering were studied. It was found that decentering correlated with verbal mental age. Children did not begin to decenter until after MA 4:11, it was not until after MA 5:11 that essentially all children decentered. Analysis of errors and of repetitions of the sentences used showed that ability to decenter did not depend on linguistic form, although some forms (e.g. the perfect tenses) led to more incorrect responses. Decentering which involved coordination of temporal viewpoints proved the most difficult.

In a longitudinal study of the emergence of reference to time during the acquisition of language by two children (Cromer, 1971), an interesting series of findings emerged which seemed to indicate that cognitive factors placed some limitations on the grammatical forms and even the lexical elements which these children could utter. Although the study was concerned specifically with temporal reference, it appeared that the cognitive constraint was a broad one, perhaps best described in Piagetian terms as an initial egocentrism and an inability to 'decenter'. Piaget (Piaget & Inhelder, 1966) describes the developmental process of the attainment of the ability to decenter as 'the transition from an initial state in which everything is centered on the child's own body and actions to a "decentered" state in which his body and actions assume their objective relationships with reference to all the other objects and events registered in the universe'.

There seem, indeed, to be three similar processes during different stages of development to which Piaget applies this terminology. The first of these is decentration which takes place at the level of momentary perceptions. Secondly, there are processes of decentration during the sensory-motor period which apply to overall perceptual systems. Of these Piaget (1961) writes: 'It is by this decentration that the progressive construction of perceptual coordinates takes place: spatial directions which were judged initially with references to the position of one's own body, are decentered in order to be related to wider frames of reference. In contrast [to the preceding phase of systematic centrations], the construction of a space which contains all objects (including one's own body), the scheme of the permanent object, spatial causality, and objective temporal sequences, are the products of a global decentration which unfolds throughout the course of development of sensory-motor intelligence...It is clear that higher forms of decentration are not simple extensions of local perceptual decentrations of the kind that occur in the perception of a figure. On the contrary, it is perceptual activities which are responsible for local decentrations, at the same time giving rise by generalization to the global forms of decentration mentioned above in connection with, for example, systems of reference' (pp. 298–299).

The third use Piaget makes of this concept is to describe the advance from preoperational representation to operational structures. 'To take projective space as an

example, it is easy to show how the development of the coordination of perspectives, starting with one's own centred perspective, constitutes a form of decentration which is analogous to those mentioned above' (p. [illegible]). Elsewhere (Piaget & Inhelder, [illegible]) they write: 'Indeed, if we compare the preoperational subperiod between four and seven or eight with the subperiod of completion between seven or eight and eleven or twelve, we see the unfolding of a long, integrated process that may be characterized as a transition from subjective centering in all areas to a decentering that is at once cognitive, social and moral' (p. 128).

Studies of the child's developing ability to decenter spatially during the pre-operational period have been reported by Piaget & Inhelder ([illegible]). In what has come to be known as the three mountain problem, the child was presented with a model in which three mountains of differing colour and height were represented in solid form. A small doll was put on the display at different positions. The child's task was to imagine and to reconstruct by a process of inference the different perspectives the doll would see in the differing positions. Piaget used three methods in eliciting answers from the children studied. In the first method the child was given pieces of cardboard and asked to build the view as seen by the doll. A second method consisted of showing the child 10 pictures and having him pick out the one most representative of the view seen by the doll. In the third method the child was shown one picture. He then had to decide where the doll had to be placed in order to see the scene in the way shown in the picture. Piaget's findings were reported in terms of a series of developmental stages, but in general the first signs of the ability to decenter in this spatial task, i.e. to take account of the shifting perspective, appeared shortly after 5 years of age.

In Cromer's study of temporal reference during the acquisition of language, it had been found that various kinds of reference apparently requiring an ability to decenter temporally, i.e. to take account of varying relationships in time from different points of view in time, began to emerge at about 4½ years of age. The purpose of this experiment is to examine systematically the development of this ability in order to find out at what age the notion of temporal relativity becomes established, so as to supply evidence which is important for analysing the extent to which notions of linguistic reference proceed. To do this, two types of tests were designed. These two types represent attempts to construct temporal decentering problems which would be analogous to methods 2 and 3 in the Piagetian spatial decentering task.

Method

Nature of the task

The problem of placing a doll at another point of view in time and asking a child to see a temporal event from that viewpoint was dealt with by employing pictures telling a story of a sequence in time. In these a little boy or girl appears, and when the entire story is represented visually, the child can be asked either to identify the picture in the story in which the pictured figure could make certain utterances, or he could be asked to take the point of view of the figure in one of the pictures and identify other points in time relative to that point. These two different procedures can be designated Type I and Type II for convenience.

These tests present a number of difficulties including such problems as teaching the young child the cartoon conventions of the 'speaking bubble' which shows that the character is talking, and presentation of left to right (or any chosen order) as representing earlier to later in time. There

[illegible]

Materials and procedure

[illegible] child's head. The warm-up items used were

1. the girl playing with the doll
   the rabbit eating the carrot
2. the boy with the blue belt
   the girl with the blue hat
3. the boy is talking
   the boy is sitting in a chair
4. the dog with the black tail
   the cat with the brown face
5. the boy playing with the lorry
   the mouse eating the cheese
6. the girl is cooking
   the girl is saying 'hello'.

[illegible]

[illegible] referred to as the Farm Story and the Zoo Story. [illegible] the lower part of Plate 1. Half the children were given the Farm Story first, and half [illegible] the Zoo Story first. As in the training series, these stories were presented [illegible] as the story was told, until all the pictures were in front of the child from left to right. Again, it was

emphasized what was earlier and later in the story. The Type I stories and the questions asked, are given below.

Farm Story. Once upon a time there was a little girl who lived in a house in the city. One day she went in a car to visit a farm in the country. There she saw a cow in the field. She walked across the field and picked flowers. Then she watched some birds fly by. Later she helped paint the barn. Then she came home and told her father all she had done.

And see, the little girl is talking in all the pictures. Show me, in what picture can the little girl say: 1, I see birds (5, i.e. fifth picture in bottom of Plate 1 is correct). 2, I'm painting the barn (6). 3, I'm going to ride to the farm in a car (1 or 2 allowed). 4, I painted the barn (7). 5, I'm picking flowers (4). 6, I saw birds (6 or 7). 7, I was painting the barn (7). 8, I will see a cow (1 or 2). 9, I have seen birds (6 or 7). 10, I will have picked flowers (1, 2 or 3).

Zoo Story. Once upon a time there was a little boy who lived in a house in the city. One day he went out walking. He went to see the zoo. There he saw a lion in a cage. Then he saw a tall giraffe. Later he saw an elephant lifting a big stick. Then he ran home and told his father all he had seen. And see, the little boy is talking in all the pictures. Show me, in which picture can the little boy say: 1, I see a lion (4). 2, I'm watching an elephant lift a big stick (6). 3, I'm going to go out walking (1 or 2 allowed). 4, I watched an elephant lift a big stick (7). 5, See me walking to the zoo? (2 or 3). 6, I saw a tall giraffe (6 or 7). 7, I was watching an elephant lifting a big stick (7). 8, I will see a lion (1, 2 or 3). 9, I have seen a giraffe (6 or 7). 10, I will have seen a lion (1, 2 and 3).

In the questions to the two stories, it can be observed that the following forms were used: 1, present; 2, present progressive (protraction or spreading events out in present time); 3, future of intent; 4, past (weak form: -ed); 5, present progressive (protraction or spreading events out in present time); 6, past (strong form); 7, past progressive (protraction or spreading events out in past time); 8, future (in one story with 'will', in the other with ''ll'); 9, present perfect (past time related to the present); 10, future perfect.

It can be seen that this Type I task is analogous to Piaget's third method in the three-mountain problem, spatial decentering test. In that method, Piaget showed the child one picture; the child's task was to decide where the doll would have to be in order to see the scene in the way shown in the picture. Analogously, in this Type I procedure the child is presented with a verbal utterance and he must show where the boy or girl in the story would have to be in the temporal sequence in order to make such an utterance.

Training items and test: Type II. The training series for the Type II test consisted of the same three three-picture sequences, with one difference: In the pictures in these sequences, the 'speaking bubble' appeared only in the middle picture. An example of another of these three training sequences, this time as used for training for the Type II test, appears in the upper portion of Plate 2. Again the pictures were placed in front of the child one by one, as the following was told:

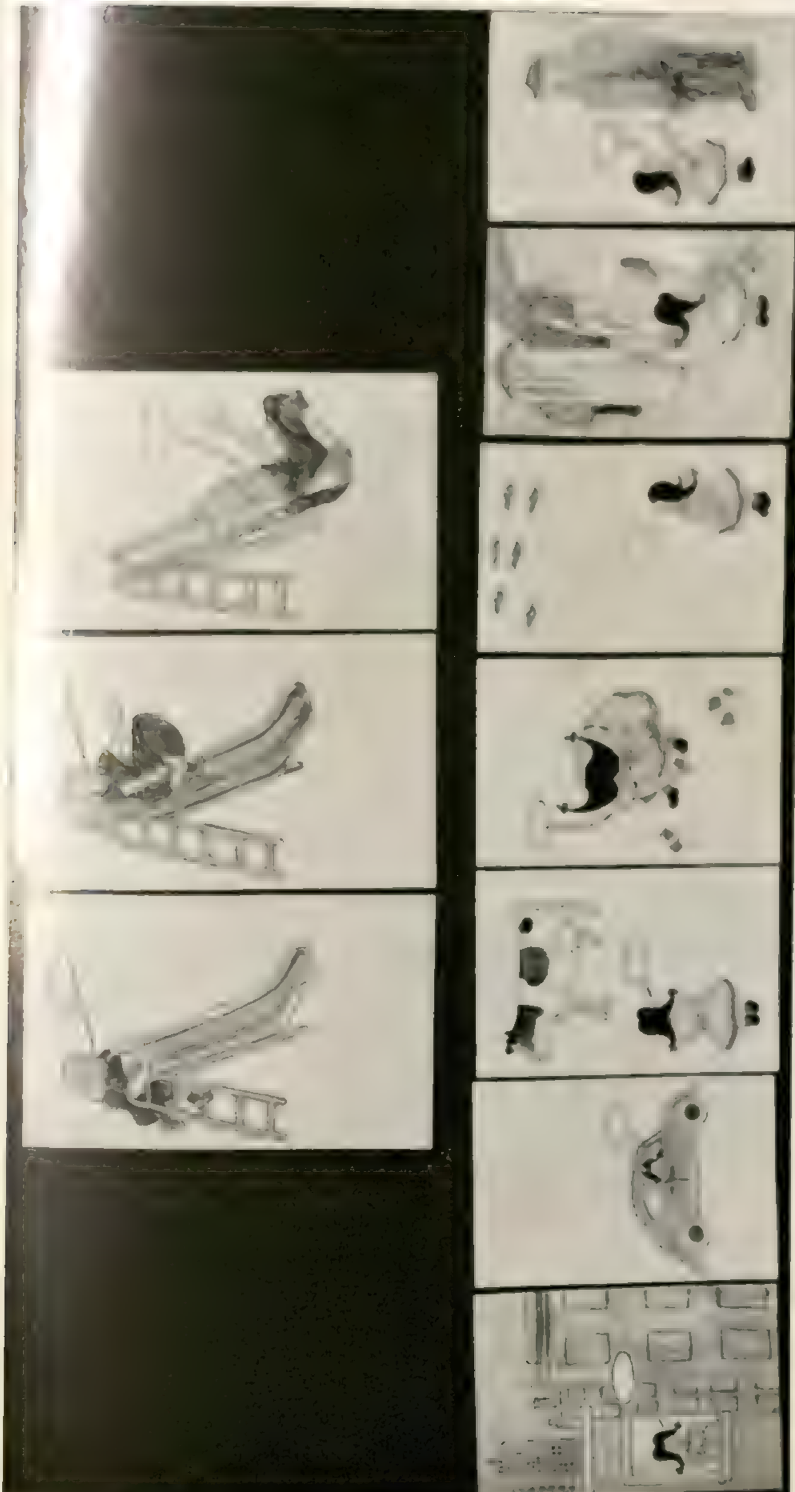
'Remember this story? "The little boy bought a balloon from a man. The little boy walked home with his balloon. The balloon burst." But see, this time he's only talking here [indicating middle picture] while he's walking home with his balloon. Show me, which picture is he talking *about*, when he says here [indicating middle picture again]: "My balloon will break", "I bought a balloon"'.

The left-to-right convention was again stressed as representing earlier to later, and the correct answers were carefully explained.

The child was then given the two Type II stories. These can be referred to as the Circus Story and the School Story. The Circus Story is shown in the lower half of Plate 2. In the two stories 'speaking bubbles' appear only in the middle and extreme end pictures.

Half the children received the Circus Story first and half the School Story first. As in the other test situation, the pictures were presented one by one as the story was told until all the pictures were in front of the child from left to right. The two Type II test stories are presented together with the questions asked, below:

Circus Story. There once was a little girl who wanted to go to the circus. She talked to her father about it. Then she went to the circus. There she saw some clowns. Next some horses came out and ran around the ring. Then she talked to the little girl in the seat next to her. Later she saw some more horses that did a dance. Then some more clowns came out and made her laugh. Later she went home and told her father all she had seen. And see – remember when she talked to the little girl in the seat next to her? [Indicate middle picture.] Show me, which picture is she



R. F. CROMER

(Facing p. 356)

R. F. CROMER

talking *about* when she says here [again indicating the middle picture]: 1. I will see some clowns (6 is the correct answer). 2. I told my father, 'I'm going to the circus' (1). 3. I'm going to see some horses (5). 4. I watched some horses (3). 5. I'll tell my father 'I went to the circus' (7). 6. I saw some clowns (2). 7. I have seen some horses (3). 8. I will have watched some clowns (6).

School Story. Once upon a time there was a little boy. He woke up one day, talked to his mother, said goodbye, and went to school. There he played ball. Later he had a glass of milk and some biscuits to eat. Then, after rest period, he talked about everything. Later, he had lunch. Then he played ring-around-the-rosy. After school he came home and told his mother all he had done.

And see—here's the little boy after rest period [indicating middle picture] when he talked about everything. What picture is he talking *about* when he says here [again indicating middle picture]: 1. I told my mother I would go to school (1). 2. I'll eat some food (5). 3. I played at school (2). 4. I'm going to play at school (6). 5. I ate some food (3). 6. I will tell my mother I went to school (7). 7. I have played at school (2). 8. I will have eaten some food (5).

It can be seen that the questions to the two stories are concerned with the same type of temporal reference, but are asked in a slightly different order in each story. Thus they concern: future ('will' in one story; ''ll' in the other); future embedded in past time; future ('going to'); past (weak form: -ed); past embedded in future time; past (strong form); present perfect (past time related to the present); future perfect.

It can also be seen that the embedding in the Circus Story is done by direct quotation, and in the School Story by indirect quotation.

This Type II test is similar to Piaget's second method in which the child was shown 10 pictures of the three mountains and had to pick out the one most suited to the view seen by the doll. In the Type II test the child is shown pictures with identical elements on either side of a central point. His task is to pick the one most suited to the viewpoint of the pictured character as expressed in some utterance.

Subjects

The subjects were 70 children at a nursery school and at an infant school in central London. The chronological age range was 3:11 to 7:4 at the time of the study. No child was used whose verbal IQ as assessed by the Peabody Picture Vocabulary Test (PPVT) was below 60, or whose parents were not native English speakers.

Other aspects of the procedure

Each child was tested individually in a session lasting about 30 min. The session began with the administration of the PPVT to establish a verbal mental age. This was done to establish a relative ordering of developmental level, based on verbal materials, in children of similar chronological age. The PPVT was used since it provides such an ordering based on lexical and not on grammatical grounds. Following this, the child was given the warm-up items, and then the training series and test for both types of decentering as described previously.

In order to ascertain whether there was any relation between failure to decenter and acoustic or memory difficulties associated with particular sentences, several days later the children at one of the schools (41 of the 70 children) were presented with sentences for immediate recall. The child's task was to repeat the sentences read to him by the experimenter, exactly as he heard them. The sentences read were the ones used in the decentering experiment, and each child heard these sentences in the same order as he originally heard them when they were used as the instructions on the decentering series.

RESULTS

Type I decentering

Responses to the questions can be classified into two basic types. A response can be labelled 'content-coded' when the child points to a picture which encodes features mentioned in the questioning sentence. For example, pointing to the picture with birds in it would be the content-coded response for such statements as 'I see birds',

'I saw birds' and 'I will see birds'. For the first of these examples, the content-coded response is correct. The latter two, however, require the second type of response – a decentered response – to any picture following the one with birds in it for 'I saw birds', and to any picture prior to the one with birds in it for 'I will see birds'. It should also be noted that decentered responses can be correctly or incorrectly decentered.

Table 1. *Percentage of children at each of four verbal mental-age groups making decentered responses on Type I stories (Farm and Zoo)*

Groups by verbal mental age on PPVT (years:months)	Number of decentered responses (maximum = 12)		
	0	1	2 or more
Group 1 (MA 3:1–4:0), $n = 12$	83·33 %	16·67 %	0·00 %
Group 2 (MA 4:1–4:10), $n = 23$	69·57 %	26·09 %	4·35 %
Group 3 (MA 4:11–5:9), $n = 22$	40·91 %	9·09 %	50·00 %
Group 4 (MA above 5:11), $n = 13$	7·69 %	7·69 %	84·62 %

$S = 879$, $\tau_c = 0{\cdot}538$, $z = 5{\cdot}13$, $P < 0{\cdot}001$.

Of the 10 questions on each of the two Type I stories, the Farm and the Zoo stories, there were six to which decentered responses were required – nos. 4, 6, 7, 8, 9 and 10. On the other questions, children almost invariably picked the content-coded picture which was the correct response. Thus, of 560 replies to non-decentering questions, only four were not content-coded. (Eight additional non-content-coded responses were by older children who decentered to the end picture when decentering was not appropriate.) However, younger children (in terms of verbal mental age) continued to choose the content-coded pictures for questions requiring decentering, whereas older children gave decentered responses. Furthermore, the number of decentered responses increased with mental age. There was a significant correlation between mental age and total number of decentered responses on the two stories ($r = 0{\cdot}57$; $n = 70$; $P < 0{\cdot}001$) However, it should also be noted that as the sample was restricted in chronological age (all but four children being between 4:2 and 6:11 years of age), there was also a significant correlation between IQ and number of decentered responses ($r = 0{\cdot}45$; $n = 70$; $P < 0{\cdot}001$). Nevertheless, a significant correlation obtained between mental age and decentering when IQ was partialled out ($r_{12\cdot3} = 0{\cdot}38$; $n = 70$; $P < 0{\cdot}01$).

The number of decentered responses does not rise smoothly with increasing mental age, however, and the 70 children can be divided into four groups based on these verbal mental ages for a more revealing analysis. Table 1 shows the percentage of children in each mental age range who did not decenter at all, decentered once, and decentered twice or more on the 12 questions requiring decentering on the two stories. It can be seen that, in the main, the youngest children did not decenter at all. Some children in group 2 began to make tentative attempts at decentering by decentering once. Group 3 children were intermediate, about half decentering not at all, but the other half decentering more than twice each. Almost all group 4 children decentered, and did so regularly on the responses requiring that ability. These trends in Table 1 were significant beyond the 0·001 level ($S = 879$; $\tau_c = 0{\cdot}538$; $z = 5{\cdot}13$).

Table 2 shows the mean number of decentered responses for each of the four groups out of a possible total of six responses requiring decentering on each of the two Type I stories. An analysis of variance revealed that there was no difference between the two

stories ($F = 0{\cdot}52$, n.s.), and no interaction effects ($F = 0{\cdot}30$, n.s.). Differences between the four groups were significant however ($F = 17{\cdot}43$, d.f. = 3, 66, $P < 0{\cdot}001$). Individual comparisons of the groups by the Scheffé method revealed that the means of groups 1 and 2 did not differ ($t = 0{\cdot}23$, n.s.), but group 3 significantly differed from these two groups ($t = 6{\cdot}14$; $P < 0{\cdot}001$), as did group 4 ($t = 9{\cdot}22$, $P < 0{\cdot}001$). The means of groups 3 and 4 also differed significantly from one another ($t = 3{\cdot}81$, $P < 0{\cdot}01$).

Table 2. *Mean number of decentered responses on Type I stories, for the four verbal mental-age groups*

Verbal mental age groups on PPVT	Story: Farm (six possible decentered responses)	Story: Zoo (six possible decentered responses)	Total (both stories) (12 possible decentered responses)
Group 1 (MA 3:1–4:0)	0·08	0·08	0·16
Group 2 (MA 4:1–4:10)	0·17	0·17	0·34
Group 3 (MA 4:11–5:9)	2·18	1·86	4·04
Group 4 (MA above 5:11)	3·54	3·46	7·00

Some of the responses of older children were decentered but incorrect. An analysis based only on correctly decentered responses yielded significance levels identical to those above with one exception: the difference between groups 3 and 4 was more highly significant ($t = 4{\cdot}46$; $P < 0{\cdot}001$), group 3 children who decentered more often getting their answers wrong than the older (group 4) children who decentered.

Type II decentering

In the Circus and School stories, for each question there were two content-coded pictures, one on each side of the story's mid-point where the pictured character spoke. The task was to decenter to the proper picture. Thus, in contrast to Type I stories, decentered responses in this task consisted of pointing to a content-coded picture – and specifically to the correct content-coded picture. It becomes evident, then, that as long as children point to content-coded pictures, random choice or guessing should lead to about 50 per cent correct responses, or four of the eight questions on each story. It should be noted that children, in fact, did choose the content-coded pictures and did not simply point anywhere in the series. For example, for the direction, 'show me, which picture is the little girl talking about when she says "I will see some clowns"', children pointed to the one of the two pictures of clowns even for incorrect responses. Of 1120 responses by the 70 children, only 10 were not content-coded, and nine of these were by two older children who often pointed to the final picture in the series, perhaps attempting to decenter in the Type I manner; such responses were counted wrong.

Table 3 shows the mean number of correct responses by each of the four mental-age groups on each of the Type II stories, Circus and School. An analysis of variance revealed that, as in the Type I stories, there was no significant difference between the responses to the two stories ($F = 1{\cdot}16$; d.f. = 1, 66; n.s.); and no interaction effects

($F = 0{\cdot}62$; d.f. $= 3, 66$; n.s.). However, differences between groups were significant ($F = 3{\cdot}79$; d.f. $= 3, 66$; $P < 0{\cdot}05$). Using the Scheffé method to compare the group means individually revealed that groups 2 and 3 did not differ from one another ($t = 0{\cdot}15$, n.s.), nor did they differ significantly from the youngest children, group 1 ($t = 2{\cdot}35$, n.s.). The oldest children, group 4, however, differed significantly from group 1 ($t = 4{\cdot}19$; $P < 0{\cdot}01$), and from the intermediate groups 2 and 3 ($t = 2{\cdot}90$; $P < 0{\cdot}05$).

Table 3. *Mean number of correct responses on Type II stories for the four verbal mental-age groups*

Verbal mental-age groups on PPVT	Story: Circus (eight possible correct answers)	Story: School (eight possible correct answers)	Total (both stories) (16 possible correct answers)
Group 1 (MA 3:1–4:0)	4·17	3·50	7·67
Group 2 (MA 4:1–4:10)	4·57	4·57	9·14
Group 3 (MA 4:11–5:9)	4·55	4·68	9·23
Group 4 (MA above 5:11)	5·77	5·23	11·00

The question can also be posed as to whether or not the mean number of correctly decentered responses on Type II stories for each group differed significantly from chance guessing. With 16 responses, guessing should produce a mean of about eight correct responses. The means of the youngest group and the two intermediate groups (7·67, 9·14 and 9·23) did not differ significantly from chance ($t = 0{\cdot}52$, d.f. $= 11$; $t = 1{\cdot}96$, d.f. $= 22$; and $t = 1{\cdot}87$, d.f. $= 21$). However, the mean for the oldest children (11·00) did differ significantly from chance guessing ($t = 3{\cdot}82$; d.f. $= 12$; $P < 0{\cdot}01$).

Two of the questions on each story used the perfect tenses (e.g. 'I have seen some horses') and it will be seen in the error analysis below that even the oldest children do less well on these. It may be that some groups are not performing at chance level on forms which they know, but that guessing on the difficult forms affects their overall scores enough so that their overall scores do not differ significantly from chance guessing. Therefore an analysis was made on answers to 12 rather than 16 questions, eliminating the two questions using perfect tenses from both stories. When this was done, the mean for group 1 (6·16) still did not differ from chance guessing, based in this case on 6 of 12 possible responses correct ($t = 0{\cdot}24$; d.f. $= 11$; n.s.). However, the means for groups 2 and 3, 7·30 and 7·41, did differ significantly from chance ($t = 2{\cdot}56$, d.f. $= 22$, $P < 0{\cdot}01$; $t = 2{\cdot}73$, d.f. $= 21$, $P < 0{\cdot}01$); and the mean for group 4, 9·00 correct responses, differed from chance at a highly significant level in this analysis ($t = 5{\cdot}01$; d.f. $= 12$; $P < 0{\cdot}001$).

Order effects

Whether a child received Type I decentering or Type II decentering stories first did not affect the number he decentered on Type I ($t = 0{\cdot}15$; d.f. $= 68$; n.s.) nor did it affect the number correct on Type II decentering ($t = 0{\cdot}04$; d.f. $= 68$; n.s.).

Error analysis

Type I. Group 1 children had only two decentered responses out of a total of 144 possible. One of these was on the past progressive form by one child, and the other was an incorrectly decentered response to the future perfect by another child, this latter probably being a random response. In group 2, of 276 possible decentered (non-content-coded) responses (23 children each having 12 possible decentered responses) there were only eight actual decenterings. Four of these were to the weak past (-ed) of which two were incorrectly decentered. There were three decentered responses to the past progressive and one to the strong past.

Table 4. *Percentage correct of decentered responses on Type I stories by groups 3 and 4 on the various linguistic forms*

	Mental-age groups on PPVT	
Linguistic forms	Group 3 (MA 4:11–5:9)	Group 4 (MA above 5:11)
-ed past	94·44 % (17/18)	86·67 % (13/15)
Strong past	83·33 % (10/12)	87·50 % (14/16)
Past progressive	89·47 % (17/19)	100·00 % (16/16)
Perfect	93·33 % (14/15)	100·00 % (17/17)
Future ('ll and will)	33·33 % (4/12)	88·89 % (8/9)
Future perfect	30·77 % (4/13)	38·89 % (7/18)

In group 3 some children decentered several times and some did not decenter at all. When children in this age range decentered, they did so on all forms about equally well with the exception that there was a slight tendency not to decenter as much on the future form using ''ll'. When a child does decenter on a particular linguistic form, it is possible to ask whether he is able to do so correctly. Table 4 shows the percentage correct of those decentered, and it can be seen that when children decentered they did so mainly correctly (over 83·3 per cent) except for the two forms, future (either 'will' or ''ll') and future perfect – both less than one-third correct.

Group 4 children – those children over verbal mental age 5:11 – almost all decentered, and they did so often. They did so about equally on all forms except the future using ''ll' in the Farm Story. Furthermore, when they decentered they decentered correctly except on the complicated future perfect form, e.g. 'I will have seen a lion' (see Table 4). On this form, just over one-third were able to decenter correctly, while on all other forms over 86·67 per cent decentered correctly.

Type II. Given two content-coded pictures, one on each side of the mid-point of the Type II stories, guessing could lead to about half of the children in any group getting the right answer on any particular form. It is thus possible to use z values for various significance levels and calculate confidence intervals for each group to see if any of the temporal forms are answered correctly significantly more often than chance. Those which reached the 0·05 and 0·01 levels are shown in Table 5 for each of the four groups. Thus, the youngest group decentered correctly significantly more often than chance on two of the complicated embedded forms. With 16 forms examined in the two stories (eight in each) one might expect one or two forms to reach the 0·05 level of significance simply by chance. But these two forms, along with the other two embedded forms, were also correctly decentered significantly more often than chance

by the other groups. Indeed, these embedded time structures were the only forms which reached even the 0·05 level in any of the groups including the oldest children, although it should also be noted that all forms except those involving the perfect tenses approached the 0·05 level in group 4.

Table 5. *Forms correctly decentered significantly more often than chance (P < 0·05) on Type II stories by the four mental age range groups*

Mental-age range group on PPVT	Story: Circus	Story: School
Group 1 (MA 3:1–4:0)	Future embedded in past	Past embedded in future
Group 2 (MA 4:1–4:10)	Future embedded in past Past embedded in future*	Future embedded in past* Past embedded in future
Group 3 (MA 4:11–5:9)	Past embedded in future*	Future embedded in past* Past embedded in future*
Group 4 (MA above 5:11)	Future embedded in past* Past embedded in future	Future embedded in past Past embedded in future

The specific sentences for these forms are:

Circus Future embedded in past: 'I told my father, "I'm going to the circus"'.
Past embedded in future: 'I'll tell my father, "I went to the circus"'.

School Future embedded in past: 'I told my mother I would go to school.'
Past embedded in future: 'I will tell my mother I went to school.'

* P < 0·01

Repetition data

When 41 of the 70 children repeated sentences read to them under the instructions to do so exactly as they heard them, most children repeated the sentences back without any changes, regardless of whether they had decentered or not, or whether their decentered responses had been correct or not a few days earlier. Children who changed words or forms of words on repetition belonged about equally to those who on the test had not decentered, decentered incorrectly, or correctly decentered for the very sentences they had changed. Thus, there was no relation between ability to repeat back the sentences exactly and ability to decenter.

One form is of special interest. It might be thought that the future using the contraction ''ll' would be missed more often than others in that the contraction would be difficult to hear and often omitted. In the Farm Story, for the sentence 'I'll see a cow', of the 36 (of the 41) children who did not decenter on this statement, 20 repeated the entire sentence back exactly as given, only five omitted the ''ll', and 11 actually expanded the form into 'I will' – even though failing to decenter. If one considers only the oldest group, in which all but one child can be said to have the ability to decenter, of those who did not decenter on this form, only one dropped the ''ll'. Seven gave perfect repetitions, and one child expanded the sentence to 'I will'. The one child who had decentered, but incorrectly, repeated the sentence back as 'I saw the cow', and the two children who correctly decentered repeated the sentence back perfectly. Similarly, in the School Story, for 'I'll eat some food', of the 13 children in the oldest group, all but one of whom decentered several times for other forms, of the five who did not decenter for this instruction, three repeated the sentence back exactly, and two expanded the ''ll' to 'I will'; none omitted the contraction.

Discussion

The results from this study on the child's ability to decenter temporally show several stages. At the youngest ages, verbal mental age 3:1 to 4:0, children did not decenter at all. At a second stage, however, at mental ages 4:1 to 4:10, children began to perform intermediately on Type II decentering tasks. Furthermore, some of these group 2 children had shown a tentative ability to decenter on Type I tasks, decentering on one response but reverting back to non-decentered responses for all other instructions. 26·09 per cent of group 2 children performed in this way. Only one child decentered more than once on the two Type I stories.

Table 6. *Summary table of performance on the two types of decentering tasks by the four groups of children*

Mental-age groups	Task: Type I	Task: Type II
Group 1 (MA 3:1–4:0)	Fail to decenter	Fail to decenter
Group 2 (MA 4:1–4:10)	Fail to decenter	Intermediate performance
Group 3 (MA 4:11–5:9)	Some of them decenter, some do not	Intermediate performance
Group 4 (MA above 5:11)	Essentially all decenter	High performance

Group 3 children, mental ages 4:11 to 5:9, were intermediate on the Type II stories with scores nearly identical to scores of group 2 children, but they had moved ahead on Type I tasks, 50 per cent of them decentering more than twice. Group 4, the oldest children, with mental ages above 5:11, were able to perform both types of decentering tasks. Eighty-four per cent decentered more than twice on Type I decentering tasks, and the mean was 7·00 decentered responses on the two stories. Similarly, they correctly decentered on Type II sequences and had significantly more correct responses on Type II tasks than did the intermediate groups. These results are represented in Table 6.

Decentering did not seem to depend on the specific linguistic forms of the verbal directions, although when decentering occurred, some difficult forms did give trouble. Thus even the oldest children decentered less often on the perfect tenses, and intermediate children decentered less often on both the future and the perfect forms. Furthermore, the repetition data showed that children who did not decenter could nevertheless perfectly repeat the sentences. It is therefore unlikely that the failure to decenter was due either to misperception or misremembering of the sentences. Since children could repeat the sentences (a linguistic task) even when they did not decenter, there is evidence that the cognitive ability to decenter develops independently of specific linguistic ability to imitate particular forms. However, children who had the ability to decenter would sometimes show difficulty with more complicated linguistic forms. They would still decenter, but do so incorrectly (see Table 4).

The Type II decentering stories both contained rather complicated forms involving embedding the past in future time, and embedding the future in past time. Children of all ages gave more correct answers to these forms than to simpler forms, and indeed, only these complicated forms reached levels of correct answers significantly above

chance. It might be possible to argue that these forms did not cause trouble because the children could have dealt with them in a cognitively simple manner which nevertheless would have yielded the correct answer. For example, all four sentences involving embedding could be shortened in the following way: 'I told my mother I would go to school' could simply be processed as 'I told my mother' and the child would still point correctly to the first picture in the series. Thus, the child would give the correct answer without dealing with the complicated temporal relations. Such an interpretation is not tenable, however, since children did not perform correctly on those instructions using the simple past (either strong or weak forms) or the simple future. How, then, could some of even the youngest children perform correctly on these complicated forms but not on the simple forms?

One possibility is that these end pictures, containing persons to whom speech could be addressed, are somehow simpler for that reason. This would not explain, however, why children are more likely than chance to choose the *correct* picture of the two, since both are 'content-coded' and contain an addressed person. Furthermore, in the Type I stories, only the seventh picture contains a person to whom speech could be addressed, and this fact did not influence the results on those tests.

A more plausible explanation is possible. It can be noted that in the Type II stories the two alternative content-coded pictures from which the child must make a choice are temporally quite complex except for the two end pictures (see Plate 2). For example, the correct answer for 'I watched some horses' is the third picture in the series. In terms of temporal relationships, that picture is 'future' with respect to pictures one and two but 'past' with respect to all the other pictures of the series. The child is asked to take the viewpoint of the middle picture. For this instruction, then, he is asked to choose, as a picture coding past time, one which is past to some viewpoints but in the future to other viewpoints. The end pictures do not share this complication and these end pictures are the content-coded representations for all four embedded forms. The correct answer for 'I told my father, "I'm going to the circus"' is the first picture in the series, and this first picture has the same relation to all other pictures in the series: it is past time with respect to all of them. Whether the child is processing a complicated embedded temporal relationship or whether he is only using a simple portion of it (e.g. 'I told my father') which happens to yield this correct answer cannot be determined. But choosing a picture representing past or future time is simpler for any form represented by content-coded pictures at the ends of the series. Here again, then, one finds a temporal analogue to Piaget's perceptual and spatial studies. In discussing the problem of seriation, for example, Piaget (Inhelder & Piaget, 1959) notes that the failure to order a series of lines at the pre-operational stage is due to the inability at that stage to coordinate two directions. Such an operation, he claims, requires reversibility – in this case the bearing in mind that a given element is both longer than those already in the series and shorter than the ones yet to follow. Similarly, in the Type II decentering experiment, all temporal choices in the series, except for the ends, require an analogous cognitive coordination.

As regards the ages at which decentering occurs it is interesting to note that in the studies of spatial decentering Piaget (Piaget & Inhelder, 1948) had found that children of CA 7 and 8 years were able to decenter, but that they still lacked the ability to coordinate viewpoints in the three-mountain problem at that age. Younger children

age 5 and 6 also showed initial ability to decenter. In the present experiment on the ability to decenter temporally it was found that on the Type I task some of the children began to decenter initially after verbal mental age 4:11, but it was not until after MA 5:11 that essentially all children decentered. The Type II task showed intermediate performance by children in the age range of 4:11 to 5:9 (groups 2 and 3), and over-all high performance by children above MA 5:11 (group 4). In the younger groups, and even in the oldest group, children gave correct answers significantly above chance guessing only on those forms which required decentering without a 'coordination of viewpoints', i.e. only to those instructions requiring decentering to the end pictures of the series – those which have the same temporal relationship to all other pictures in the series, and thus requiring no complex coordination. The oldest group (above MA 5:11) gave correct answers approaching significant differences from chance levels on the instructions requiring such coordination.

The materials used in this experiment were originally produced for related experiments by funds from Public Service Research Grant MH-7088, National Institutes of Mental Health.

I wish to thank Mrs Tebbutt, headmistress of St Leonard's Nursery, and Miss T. R. Gardner, acting headmistress of St George the Martyr School, for their kind cooperation in providing subjects and facilities.

References

Cromer, R. F. (1971). The development of temporal reference during the acquisition of language. T. G. Bever & W. Weksel (eds.), *The Structure and Psychology of Language*, vol. 2. New York: Holt, Rinehart & Winston.

Fraser, C., Bellugi, U. & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *J. verb. Learn. verb. Behav.* **2**, 121–135.

Inhelder, B. & Piaget, J. (1959). *The Early Growth of Logic in the Child*. London: Routledge & Kegan Paul, 1964.

Piaget, J. (1961). *The Mechanisms of Perception*. London: Routledge & Kegan Paul, 1969.

Piaget, J. & Inhelder, B. (1948). *The Child's Conception of Space*. London: Routledge & Kegan Paul, 1963.

Piaget, J. & Inhelder, B. (1966). *The Psychology of the Child*. London: Routledge & Kegan Paul, 1969.

(*Manuscript received* 17 *March* 1970; *revised manuscript received* 14 *August* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 367–373
Printed in Great Britain

BEHAVIOURAL CONTRAST AND PEAK SHIFT IN CHILDREN

By JOHN N. NICHOLSON AND JEFFREY A. GRAY
Department of Experimental Psychology, University of Oxford

The phenomena of behavioural contrast and peak shift, previously reported in operant conditioning studies using pigeons and rats, were obtained with school children after training on an angular-orientation discrimination task.

Research on the course of discrimination learning in the pigeon using operant conditioning techniques has repeatedly demonstrated the phenomena known as behavioural contrast and peak shift (Terrace, 1966*a*; Reynolds, 1961; Hanson, 1959). Both these phenomena are seen when subjects are first trained to make a response in the presence of a stimulus (S^D) signalling availability of reward, and then given discrimination training between this stimulus and another one (S^Δ) signalling non-availability of reward. Behavioural contrast consists in a rise in response rate in the presence of S^D during discrimination training relative to the rate obtaining before discrimination training has begun. When S^D and S^Δ are two values along a single stimulus continuum (e.g. wavelength of light), peak shift is observed after discrimination training: whereas generalization testing over the whole stimulus continuum conducted before discrimination training reveals a peak response rate at S^D, similar generalization testing conducted after discrimination training reveals a shift in the peak response rate from S^D in the direction away from S^Δ. Both behavioural contrast and peak shift may be regarded as reflecting an emotional state ('frustration') arising in response to 'frustrative non-reward' (Amsel, 1962), and a mathematical model developed by Gray & Smith (1969) is able to account for them in these terms.

If behavioural contrast and peak shift reflect basic emotional responses, it might be expected that they would be found at many levels of the phylogenetic scale. Other frustration-related phenomena have been observed in pigeons (Staddon & Innis, 1966), rats (Amsel & Roussel, 1952), monkeys (Davenport & Thompson, 1965) and children (Ryan & Watson, 1968). Behavioural contrast has so far been reported in pigeons (Reynolds, 1961) and retardate children (O'Brien, 1968), while peak shift has been reported in the pigeon (Hanson, 1959) and in the rat (Schwartzbaum *et al.*, 1964; Pierrel & Sherman, 1960, 1962). Landau (1968) has looked for peak shift in normal children, but failed to find it. La Berge (1961) and Cross & Lane (1962) report a phenomenon which is very similar to peak shift with adult human subjects. However, in both these experiments, as in Landau's, the subject was required to give one of two responses to each stimulus, S^D and S^Δ. In such a choice-response task the *rate* of emission of each response does not vary, only its *probability* of emission on each trial. It cannot be taken for granted that these two methods of measurement measure the same thing. Indeed, the separation of response probability and response vigour, which is a central feature of the Gray & Smith (1969) model, suggests that they do not.

In the present paper, we demonstrate both behavioural contrast and peak shift in normal children.

METHOD

Subjects

Subjects were rural children of average intelligence, ranging in age from 7:1 yr. to 10:6 yr. Testing was conducted at their school, Deddington Primary, Oxfordshire, during normal school hours. There were seven subjects in Expt. I (three boys, four girls); three of the girls who were slow to learn the discrimination were excluded from Expt. II.

Apparatus

The children, who were tested singly, sat in front of an aluminium cabinet within easy reach of a shaped lever and a tray. Pressing the lever sometimes resulted in the delivery of poker chips into the tray by a Gerbrands Universal Dispenser. These acted as token rewards in Expt I. In Expt. II every fourth poker chip was replaced by a sweet. Schedules of reinforcement were controlled by a film-strip timer. Lever-presses were recorded on counters and a Rustrak event recorder. The stimuli were back-projected from an automatically operated Kodak Carousel projector on to a 9 × 9 in. translucent Perspex screen in the centre of the cabinet. They consisted of coloured drawings of a space rocket orientated at various angles against a sky-blue background. The experimenter was invisible to the subject behind a partition. Instructions were kept to a minimum, emphasizing mainly that pressing the lever would sometimes result in delivery of a poker chip or (in Expt. II) a sweet and that the pictures on the screen might have some connexion with the delivery of these rewards. The children were encouraged to obtain as many poker chips as possible, although the chips were subsequently returned and not exchanged for any further reward.

Procedure

All sessions were of 30 min. No difficulty was experienced in maintaining interest and a steady response rate over this period. Total sessions per child ranged from 9 to 15.

Expt. I. Sessions 1–3 were spent in adapting the children to the situation and familiarizing them first with the response mechanism and later with the S^D (rocket in vertical position). Presentation of the S^D followed the same pattern later adopted for the discrimination training and testing: periods of 30 sec. stimulus duration separated by 1–2 sec., during which the projector switched to the next slide and the screen was blank. By the end of the third session, after a sequence of reinforcement schedules,* all subjects were responding regularly on VI 1 min., i.e. with rewards for lever-pressing becoming available, on average, once a minute.

The first generalization test was then carried out in a single session with eight test stimuli (seven stimuli whose orientation ranged in 13° steps from 51° to 129°, including the S^D, 90°, and an eighth stimulus consisting of the blue background with no rocket). These were presented without reinforcement in random order, four times each, and were interspersed with 32 additional presentations of the S^D, during which responding was reinforced on VI 1 min.

After the first generalization test, discrimination training was begun: reward continued to be available on VI 1 min. in the presence of the vertical rocket, but was never given during S^Δ, which consisted of the sky-blue background with no rocket. There were an equal number (30) of 30 sec. presentations of S^D and S^Δ, randomly intermixed, in each session. It took two to eight sessions for different subjects to reach the minimum criterion of discrimination, a 6:1 'discrimination ratio' (Dinsmoor, 1951) of response rates in S^D and S^Δ. A second generalization test, of the same form as the first, was then carried out with the four subjects who learnt the discrimination.

Expt. II. The four subjects who learnt the presence/absence discrimination of Expt. I were then trained on an intra-dimensional discrimination task, with S^D orientated 32·5° to the right of

* The schedules of continuous reinforcement (CRF), fixed ratio (FR), fixed interval (FI) and variable interval (VI) were used in a flexible fashion, as is customary in the 'shaping' of behaviour; with CRF, each response produces a reward, whereas FR, FI and VI schedules provide only intermittent reinforcement, in the first case (FR) according to a fixed proportion of responses made (one reward for every n responses), in the case of the others according to the passage of a time interval which may either be fixed (FI), when a reward will follow the first response after t sec., or variable (VI) when subjects receive, *on average,* one reward every t sec. These schedules and their use in 'shaping' are discussed by Ferster & Skinner (1957).

vertical and S^{Δ} 32·5° to the left. This discrimination was established with the same reinforcement schedule operating as in the discrimination training phase of Expt 1, in one to two sessions, using as criterion the same value of the discrimination ratio as in Expt 1. A final generalization test was then carried out, using 12 test stimuli, ranging from 18·5° to 161·5° in 13° steps, following the same procedure as in Expt. 1, but taking two sessions and involving seven presentations of each test stimulus.

Results

Expt. 1. The course of training on the discrimination between the presence and absence of the vertical rocket is shown for each of the seven subjects in Fig. 1. It can be seen that three of them (subjects 5–7) show no sign of learning the discrimination and that all of these show decreasing rates of response in both S^D and S^{Δ}. By contrast, the other four subjects (subjects 1–4) make a clear discrimination between S^D and S^{Δ}: in each of these cases the S^D response rate climbs above the pre-discrimination baseline (averaged over the last two days' pre-discrimination responding) *pari passu*

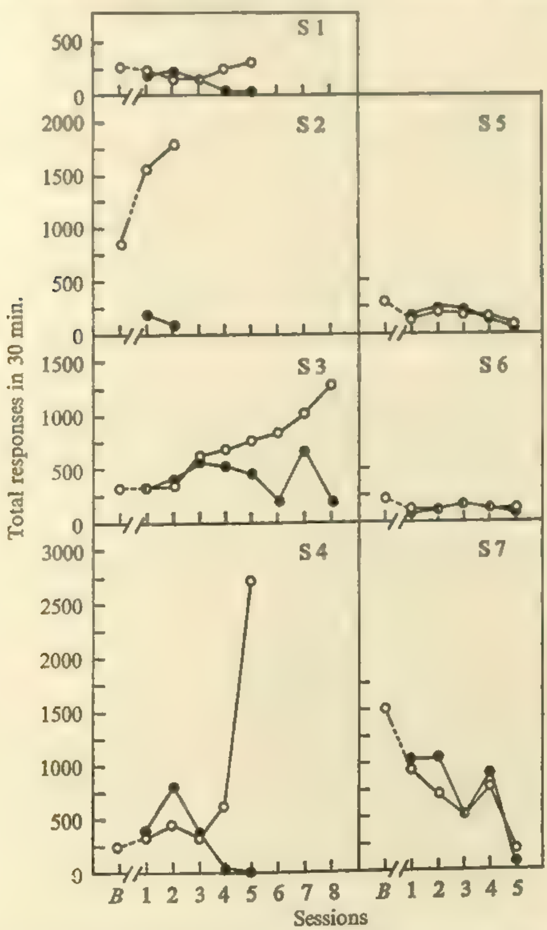


Fig. 1. Response rates during S^D and S^{Δ} for four subjects (S 1–S 4) who learnt the discrimination between presence and absence of the stimulus rocket and three subjects (S 5–S 7) who did not. The point *B* on the abscissa represents the average response rate in the last two pre-discrimination sessions in the presence of the vertical stimulus. ○——○, Responses during S^D; ●——●, responses during S^{Δ}.

with the decline in S^D response rate. By the end of discrimination training S^D response rate exceeds the pre-discrimination baseline by 125–1000 per cent: these may be regarded as measures of individual contrast effects.

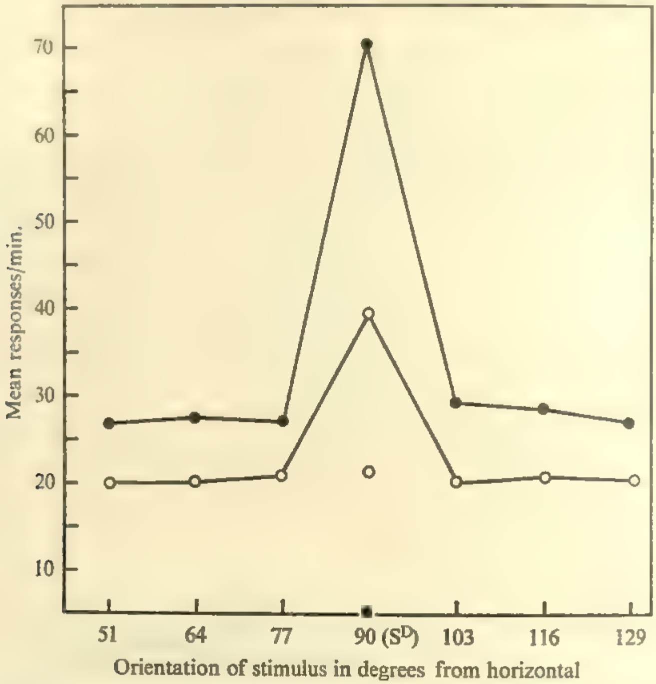


Fig. 2. Mean response rates in the presence of stimulus rockets orientated at different angles before and after training on the discrimination between presence of a vertical rocket (S^D) and its absence (S^Δ). The single points show rates during S^Δ. ●——●, After discrimination training; ○——○, before discrimination training. ●, S^Δ (after); ○, S^Δ (before).

Fig. 2 shows the results of the two generalization tests averaged over the four subjects who learnt the discrimination. It can be seen that, prior to discrimination training, they treat all orientations of the rocket other than vertical, and also absence of the rocket, as equal, responding at a rate clearly above zero but much less than the rate in S^D. After discrimination training all rocket orientations other than vertical are still treated alike, but the rate in the absence of the rocket is now virtually zero. The behavioural contrast effect is clearly evident, not only as an increase in the rate in S^D itself, but also as an elevation in the whole stimulus generalization curve. Note that the peak of this curve remains at S^D, as would be expected in the absence of intra-dimensional discrimination training.

Expt. II. The results of the final generalization test are shown for each subject and averaged over the group of four in Fig. 3. Subject 1 and subject 2 show clear evidence of a peak shift, by 13° and 26° respectively. There is no shift of the peak in subject 3 or subject 4 or in the group mean curve. A more sensitive measure of changes induced in the stimulus generalization curve by intra-dimensional discrimination training, suggested by Grusec (1968), is to treat the generalization curve as a frequency distribution and to calculate the mean of this distribution in units of the stimulus

continuum laid out along the abscissa, in this case, degrees. We may call the results of this calculation 'mean shift' in distinction to the 'mode shift' involved in the usual peak-shift calculation. (The legitimacy of the mean shift calculation will be discussed in more detail in a later paper.) The vertical lines in Fig. 3 show the results of the mean shift calculations for our data, considering S^D and the three values on either side of it. A shift is now apparent in three of the four subjects, and in the group average curve.

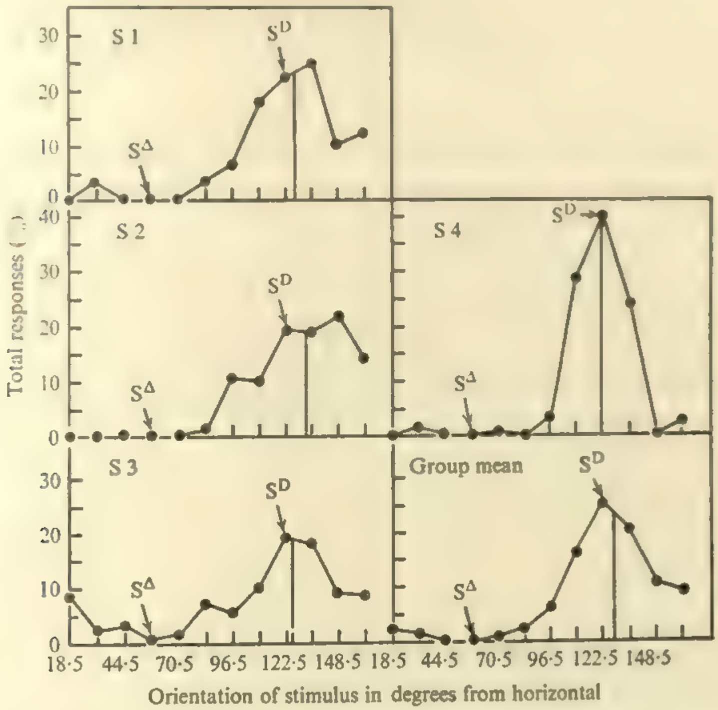


Fig. 3. Response rates in the presence of stimulus rockets orientated at different angles after training on the discrimination between 122·5° (S^D) and 57·5° (S^Δ). The vertical line shows the Grusec mean (see text).

Discussion

It would seem reasonable to conclude that children, like pigeons, may display behavioural contrast and peak shift.

With respect to behavioural contrast, however, it might be objected that we did not take sufficient care to ensure that our subjects were responding at an asymptotic rate before embarking on discrimination training. Russell (1960) has pointed out that continued exposure to a VI schedule characteristically produces a steadily increasing rate of response in S^D, regardless of the presence of another stimulus associated with a less favourable schedule of reinforcement. However, we have also found behavioural contrast in a later experiment* in which the pre-discrimination training period was sufficiently long to ensure that asymptotic response rate had been achieved.

* The results of this experiment will be described in a subsequent paper.

It should be noted that the fact that, in the present experiment, only those subjects who formed the discrimination showed behavioural contrast is entirely in agreement with the animal literature (Terrace 1966*a*).

With respect to peak shift, the failure of one subject, subject 4, to demonstrate this effect even when Grusec's (1968) measure is used is not surprising, given the extreme efficiency of his performance on the discrimination task: his S^D S^Δ discrimination ratio is 80:1, compared with ratios of from 7:1 to 40:1 in the other three subjects. Terrace (1966*a*, *b*) has shown that the peak shift in pigeons depends on the occurrence of a sufficient number of errors, i.e. responses in the presence of S^Δ: it does not appear after errorless discrimination training, and it is reduced in subjects given extended discrimination training.

The early demonstrations of peak shift (Hanson, 1959) used wavelength of light as the stimulus continuum, and it has been suggested (Guttman, 1965) that this phenomenon is the result of colour contrast. Pierrel & Sherman (1960, 1962), however, found a peak shift using an auditory intensity discrimination in rats, and the present results confirm Bloomfield's (1967) with pigeons showing that peak shift may be obtained with angular orientation. There are so many differences between our procedure and Landau's (1968), who also used angular orientation with children but failed to get peak shift, that it is difficult to be certain of the reason for his failure. However, his use of a verbal labelling choice response and/or the size of his stimulus intervals (30° compared to our 13°) may be responsible.

Our findings not only suggest continuity in some of the processes underlying the formation of visual discriminations in species as diverse as pigeon, rat and man, but also provide us with valuable tools for measuring an individual's susceptibility to frustration, since the magnitude of peak shift and behavioural contrast may be used as quantitative indices of this. Furthermore, if these phenomena can be related on the one hand to a consistent pattern of individual differences (in children) and on the other to a physiological substrate (in animals), they might provide a promising route into the study of the physiological basis of personality.

This research was supported by a grant from the Social Science Research Council of Great Britain. We are very grateful to the Headmaster and staff of Deddington Primary School, Oxfordshire, for their generous cooperation. We would also like to thank K. J. Kapota and S. J. Graham for their help in the design and construction of the apparatus.

References

Amsel, A. (1962). Frustrative nonreward in partial reinforcement and discrimination learning: some recent history and a theoretical extension. *Psychol. Rev.* **69**, 306–328.

Amsel, A. & Roussel, J. (1952). Motivational properties of frustration. I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.* **43**, 363–368.

Bloomfield, T. M. (1967). Peak-shift on a line-tilt continuum. *J. exp. Analysis Behav.* **10**, 361–366.

Cross, D. V. & Lane, H. L. (1962). On the discriminative control of concurrent responses: the relations among response frequency, latency, and topography in auditory generalization. *J. exp. Analysis Behav.* **5**, 487–496.

Davenport, J. W. & Thompson, C. I. (1965). The Amsel frustration effect in monkeys. *Psychon. Sci.* **3**, 481–482.

DINSMOOR, J. A. (1951). The effect of periodic reinforcement on barpressing in the presence of a discriminative stimulus. *J. comp. physiol. Psychol.* **44**, 354.

FERSTER, C. B. & SKINNER, B. F. (1957). *Schedules of Reinforcement.* New York: Appleton-Century-Crofts.

GRAY, J. A. & SMITH, P. T. (1969). An arousal-decision model for partial reinforcement and discrimination learning. In R. M. Gilbert & N. S. Sutherland (eds.), *Animal Discrimination Learning*. London: Academic Press.

GRUSEC, T. (1968). Peak shift in stimulus generalization: equivalent effects of errors and non-contingent shock. *J. exp. Analysis Behav.* **11**, 239–249.

GUTTMAN, N. (1965). Effects of discrimination formation on generalization. In D. I. Mostofsky (ed.), *Stimulus Generalization*. Stanford: Stanford University Press.

HANSON, H. M. (1959). Effects of discrimination training on stimulus generalization. *J. exp. Psychol.* **58**, 321–334.

LA BERGE, D. (1961). Generalization gradients in a discrimination situation. *J. exp. Psychol.* **62**, 88–94.

LANDAU, J. S. (1968). Post-discrimination generalization in human subjects of two different ages. *J. exp. Psychol.* **76**, 656–663.

O'BRIEN, F. (1968). Sequential contrast effects with human subjects. *J. exp. Analysis Behav.* **11**, 537–542.

PIERREL, R. & SHERMAN, J. G. (1960). Generalization of auditory intensity following discrimination training. *J. exp. Analysis Behav.* **3**, 313–322.

PIERREL, R. & SHERMAN, J. G. (1962). Generalization and discrimination of the S^D–S^Δ intensity difference. *J. exp. Analysis Behav.* **5**, 67–71.

REYNOLDS, G. S. (1961). Behavioural contrast. *J. exp. Analysis Behav.* **4**, 57–71.

RUSSELL, I. S. (1960). Analysis of the responding during operant discrimination. (Unpublished doctoral dissertation, Indiana University.)

RYAN, T. J. & WATSON, P. (1968). Frustrative non-reward theory applied to children's behaviour. *Psychol. Bull.* **69**, 111–125.

SCHWARTZBAUM, J. S., THOMPSON, J. B. & KELLICUTT, M. H. (1964). Auditory frequency discrimination and generalization following lesions of the amygdaloid area in rats. *J. comp. physiol. Psychol.* **57**, 257–266.

STADDON, J. E. R. & INNIS, N. K. (1966). An effect analogous to frustration on interval reinforcement schedules. *Psychon. Sci.* **4**, 287–288.

TERRACE, H. S. (1966*a*). Stimulus control. In W. K. Honig (ed.), *Operant Behavior: Areas of Research and Application*. New York: Appleton-Century-Crofts.

TERRACE, H. S. (1966*b*). Behavioural contrast and peak shift: effects of extended discrimination training. *J. exp. Analysis Behav.* **9**, 613–617.

(*Manuscript received* 6 *July* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 375–380
Printed in Great Britain

NOISE TOLERANCE AND EXTRAVERSION IN CHILDREN

By COLIN D. ELLIOTT
Hester Adrian Research Centre, University of Manchester

A method of measuring tolerance of white noise is described. The subjects were groups of children aged 5 and 10 years. Extraverts tolerated a significantly greater intensity than introverts. Boys tolerated a significantly greater intensity than girls. There was no difference between the two age groups in level of noise tolerance. The method can be used to measure extraversion in a younger age range than is possible with questionnaires. Results are discussed in terms of Eysenck's (1967) concepts of 'stimulus hunger', 'stimulus avoidance' and hedonic tone.

The personality dimension of introversion–extraversion has been interpreted by Eysenck (1955, 1957) in terms of individual differences in cortical excitation and inhibition. Evidence on the lower level of cortical arousal in extraverts has been provided in EEG studies (Savage, 1964; Gale *et al.*, 1969). It can be predicted from the theory that a given level of sensory stimulation would have a differential effect upon extraverts and introverts: in order to achieve equal subjective intensity, extraverts would require a higher level of sensory input. In support of such a prediction, it appears that extraverts have higher auditory thresholds (Smith, 1968), lower reliability of responses in audiometry (Reed & Francis, 1962) and higher pain thresholds (Haslam, 1967), while introverts show greater tolerance of very low levels of stimulation (Petrie *et al.*, 1960; Francis, 1969).

Evidence that noise has a differential arousing effect upon extraverts and introverts has been provided by Davies & Hockey (1966) and by Davies *et al.* (1969). The first of these reports indicated that high-intensity white noise improved the visual vigilance performance of extraverts significantly more than that of introverts. In the second study, varied auditory stimulation significantly reduced the number of commission errors made by extraverts but not by introverts in a visual vigilance task. It was argued that extraverts, who normally show a lower level of cortical arousal than introverts, would have their level of arousal raised to an optimal point by noise, and that this in its turn would improve their performance. Introverts are presumably much nearer the optimal level of arousal to begin with, noise consequently having little effect upon their performance.

Eysenck (1960) has suggested that individuals strive to maintain a positive hedonic tone in relation to incoming stimulation. Stimuli that are too strong or too weak produce negative hedonic tone and the individual strives to reduce or increase the level of stimulus input as appropriate. Due to a lower level of cortical excitation, extraverts will have a higher optimum level of stimulation required to produce positive hedonic tone than introverts. Thus at relatively low levels of stimulation extraverts will show 'stimulus hunger', while at relatively high levels introverts will show 'stimulus avoidance' (Eysenck, 1966, 1967). It is not entirely the case that 'introverts reduce arousal by *avoiding* stimulation and extraverts raise arousal by *seeking* stimulation' (Gale, 1969): at extreme levels of stimulation both groups would be expected to show the same characteristics.

Some evidence has been provided which lends support to the stimulus hunger hypothesis. Farley & Farley (1967) found a correlation of 0·48 between extraversion and score on a scale designed to measure sensation-seeking behaviour. Howarth (1964) found that extraverts press buttons in order to increase stimulation significantly more than do introverts. Weisen (1965), reported by Eysenck (1966), found that fast lever-pulling, in order to produce 3 sec. of bright lights and loud jazz music, was more frequent in extraverts. Davies *et al.* (1969) also found that extraverts requested 30-sec. periods of varied auditory stimulation during a visual vigilance task significantly more frequently than did introverts. In a similar experiment, Gale (1969) reported that the mean listening time of extraverts and introverts to various types of auditory stimulation produced by pressing morse keys did not differ significantly, although the group means varied in the predicted direction. Extraverts, however, obtained more stimulation by adopting the strategy of a greater frequency and variety of key-pressing. Such an increase in variety can be considered to be arousing (Berlyne, 1960). Gale could perhaps have produced significant differences between extraverts and introverts in terms of mean listening time by increasing the intensity of the auditory stimulation.

In addition to his experiment referred to above, Weisen (1965) tested the stimulus avoidance hypothesis by presenting bright lights and loud music which could be extinguished by fast lever-pulling. He found that introverts worked harder than extraverts to obtain darkness and silence. Further support for this finding is provided by Davies *et al.* (1969), who gave continuous varied auditory stimulation to extraverts and introverts during a vigilance task. Significantly more requests for periods of silence were made by introverts.

In view of the evidence cited above, there appear to be significant differences between extraverts and introverts in response to auditory stimulation. This may be interpreted as due to differential levels of arousal in the two groups. It can be predicted that, as noise is an arousing stimulus, there would be significant differences between extraverts and introverts in terms of tolerance of high levels of noise.

The previously mentioned studies employed adult subjects and did not provide separate analysis of the performance of each sex.

Although it seems likely that there are marked temperamental differences between young children (Thomas *et al.*, 1963), it has been found difficult to measure these reliably. There is little information on developmental changes in temperament: personality questionnaires have not been successful with young children. S. B. G. Eysenck (1965*a*) found, when standardizing the Junior Eysenck Personality Inventory, that 'as far as extraversion is concerned, either this is a personality dimension which is not clearly emerging until the age of 9 or 10, or its measurement by means of questionnaires presents unusual difficulty with the youngest age groups'.

Consistent differences between the sexes have previously been noted in terms of extraversion as measured by personality questionnaires, on which males tend to score more highly than females (S. B. G. Eysenck, 1965*b*). Such differences may be constitutional or may perhaps be accounted for by a greater degree of social inhibition being acquired by females than by males because of cultural demands. If noise tolerance is a more direct measure of cortical arousal than personality questionnaires,

which tend to measure social inhibition (Cattell, [illegible]), a difference between the sexes in terms of noise tolerance may strengthen the view that this has a constitutional basis.

The present experiment was designed to examine differences between extraverts and introverts of both sexes and at two age levels in terms of noise tolerance.

Method

Subjects. Sixty-four subjects were divided into two age-groups as follows: (1) older group: 16 boys and 16 girls aged between 10 years and 10 years 11 months; (2) younger group: 16 boys and 16 girls aged between 5 years and 5 years 11 months. All the children attended a junior or an infant school, depending on their age group, situated in a large county borough. Only one school was used for each age-group.

Any child in the school who was within the age-groups concerned was eligible for selection, with the following exceptions: (a) children of obviously subnormal intelligence, who would be unable to follow simple verbal instructions, (b) children known to have a hearing loss, (c) children known to the teachers for less than 3 months.

Children were rated for extraversion by their head teachers and their class teachers.

Although teachers' ratings of extraversion do not always agree with inventory responses (Eysenck & Cookson, 1969), this was the only method by which the 5-year-old children could be selected. The older children were also selected by teachers' ratings in order to maintain consistency in the experimental design. The teachers were instructed in the method of rating the children according to the technique recommended by Warburton (1962), who also gives a description of Cattell's factor of exvia–invia. This description was used as the basis of a list of characteristics of extraversion–introversion on which the children were rated. The teachers were first asked to select the five most extraverted and the five most introverted children in the age-group for each sex. The remainder of the children in the age-group were divided into fairly extraverted and fairly introverted groups, the sexes being considered separately. Five of each of these groups were selected randomly and allocated to subgroups.

From the five children in each subgroup, four were randomly selected for participation in the experiment.

To check the validity of the results, the older group of children also completed the Junior Eysenck Personality Inventory (S. B. G. Eysenck, 1965*a*).

Apparatus. A Peters Audiometer with a Bekesy automatic attachment was used.

Wide band white noise was presented binaurally through earphones. A press-button switch was connected to the Békésy attachment in such a way that when it was pressed the intensity of sound decreased, and that when it was released, the intensity of sound increased.

An automatic stop was fitted in the Bekesy attachment which prevented the sound intensity from exceeding 110 db. This was to prevent damage occurring either to the machine or to the individual.

A pen-recorder automatically graphed the variations in the level of sound intensity.

Procedure. All subjects in both age-groups were given sweep tests of hearing at an intensity of 20 db over the frequencies of 500, 1000, 2000 and 4000 c.p.s. Any child who failed this sweep test was not included in the experiment. The fifth child in the original subgroup selected by the teachers was then included in the experiment. Four children were excluded in this way.

The subject was placed on a chair facing a wall, and was then given the following instructions: 'Here are some earphones. They make a hissing noise. Listen.' The earphones were placed near the subject's ear, with the intensity of the white noise set at the level of 40 db. The experimenter then produced the press-button switch, and said: 'Here is a button. You can press it with your thumb. Now, when I put these earphones on you this hissing noise is going to get louder and louder and louder. When you don't want it to get any louder, press this button. When you press the button the noise will get softer and softer and softer. Then when the noise is soft enough, you let the button go, and the noise will get louder and louder again.' The instructions were repeated until the subject understood. He was instructed to keep performing according to the instructions until the noise stopped.

The earphones were placed on the subject and the Békésy attachment was started at the level

of 40 db. During the first minute of operation the subject was encouraged to [illegible] tions if in any difficulty, or, more likely, if he was playing with the button to [illegible].

The Bekesy attachment, set at its fast speed, takes 2 min 45 sec to run. Performance during the first minute was not scored. The mean of the six highest levels of intensity, which were recorded in the final 1¾ min., was used as a measure of noise tolerance.

In order to confound any possible effect of the time of day on the performance of the subjects (Colquhoun & Corcoran, 1964), half of each group were tested in the morning between 9.30 a.m. and 12 noon, and half in the afternoon between 1.30 p.m. and 4 p.m.

A sample of 16 children balanced for age and sex was retested with both experimental measures in order to obtain an estimate of test–retest reliability. The length of time between tests varied from 1 week to 4 weeks, depending on when the child was first tested.

RESULTS

The mean scores of the various groups, expressed in terms of the mean maximum intensity (db) of white noise tolerated by the children, are given in Table 1. It will be seen that the mean scores for the extraverted and introverted groups vary in the predicted direction. The results confirm that extraverts have a greater tolerance of noise than introverts, the former group having a mean tolerance level over 30 db higher than the latter.

Table 1. *Means and standard deviations of the various groups*

Group	Mean (db)	S.D. (db)
Highly extraverted ($n = 16$)	96·24	12·19
Fairly extraverted ($n = 16$)	76·00	18·91
Fairly introverted ($n = 16$)	72·28	14·67
Highly introverted ($n = 16$)	63·70	15·11
Age 10–11 years ($n = 32$)	77·92	16·59
Age 5–6 years ($n = 32$)	76·19	21·96
Boys ($n = 32$)	81·54	20·07
Girls ($n = 32$)	72·57	17·82
Total ($n = 64$)	77·06	19·65

Mean age differences are minimal, but there is a significant sex difference, boys having higher noise tolerance than girls.

A χ^2 test indicates that the total frequency distribution does not differ significantly from normal ($\chi^2 = 6{\cdot}76$; d.f. $= 5$; $P < 0{\cdot}05$). This is despite the fact that five children rated as highly extraverted and two rated as fairly extraverted tolerated the maximum intensity of noise (110 db) produced by the audiometer.

An analysis of variance (see Table 2) indicates that the differences between the groups in terms of extraversion are highly significant. As there are no significant interactions we can conclude that these differences are found in both sexes at both age levels. Boys also tolerate a significantly higher level of noise than girls at both age levels.

Because there was insufficient time available to retest all subjects, the test–retest reliability of this measure was calculated for 16 of the initial sample—one child randomly chosen from each cell of the analysis of variance table. The reliability coefficient was 0·854. A t test of significance gives a value of $t = 6{\cdot}14$ ($P < 0{\cdot}001$).

In order to check the validity of the teachers' ratings, a comparison was made between the mean scores for noise tolerance, in the older group only, when grouped by teachers' ratings and when grouped by scores on the Junior Eysenck Personality

Inventory. Results are given in Table 3. It should be noted that the mean scores for the extraverts and the introverts were obtained by pooling the results of the two extraverted and the two introverted groups respectively. The results indicate that the mean levels are very similar for either method of selection, the groups selected by JEPI scores being slightly more widely differentiated.

Table 2. *Summary of analysis of variance*

Source of variation	S.S.	D.F.	M.S.	F
Extraversion (E)	9123·7	3	3041·2	13·2**
Age (A)	48·3	1	48·3	< 1
Sex (S)	1287·0	1	1287·0	5·57*
E × A	1082·1	3	360·7	1·56
E × S	233·1	3	77·7	< 1
A × S	311·5	1	311·5	1·35
E × A × S	1145·2	3	381·7	1·65
Within cell	11,091·2	48	231·1	—
Total	24,324·1	63	—	—

** $P < 0{\cdot}001$. * $P < 0{\cdot}025$.

Table 3. *Comparison of mean levels of noise tolerance for groups selected by teachers' ratings and by JEPI scores*

	Teachers' ratings	JEPI scores
Extraverts	82·8	85·8
Introverts	72·1	69·1

Discussion

The results confirm the hypothesis that extraverts have a higher level of noise tolerance than introverts. The latter group show stimulus avoidance at a lower intensity of auditory stimulation than the former. The extraverts also showed stimulus avoidance, but at a higher level of stimulation. Their level of stimulus hunger was not such as to lead them all to accept the maximum available level of white noise, although seven extraverts did so with every indication of enjoyment. Furthermore, the level of stimulus avoidance in the introverts was not such as to lead any of them to completely suppress all input. This test of noise tolerance seems to fit Eysenck's (1960) theory of hedonic tone. The levels of noise tolerance, given the present instructions, probably reflect the point where there is sufficient negative hedonic tone to cause the individual to seek to reduce it. By a slight alteration of the instructions, the test could measure the level of noise which reflects maximum positive hedonic tone. It would be predicted that such a procedure would still produce significant differences between extraverts and introverts, the mean levels of noise accepted by the subjects being lower than in the present experiment.

The presence of a significant sex difference in level of noise tolerance is in agreement with previous findings of sex differences in results from questionnaires. If noise tolerance is indeed a measure of cortical arousal, the results of the present study may lend support to the view that males are significantly less highly aroused than females.

The lack of any difference between the 5- and the 10-year-olds, together with the lack of any sex × age interaction, suggests that developmental changes in level of extraversion are minimal, at least beyond the age of 5 years. The difficulties

encountered in measuring extraversion in young children with questionnaires appear to be due to the unsuitability of the technique: the present study clearly indicates that extraversion is measurable in children aged 5 years and upwards.

I am grateful to Professor I. G. Taylor for the provision of equipment for this experiment and to the teachers for the time they spent in selecting the children. The author is the holder of an SSRC Senior Research Fellowship.

REFERENCES

BERLYNE, D. E. (1960). *Conflict, Arousal and Curiosity*. New York: McGraw Hill.

CATTELL, R. B. (1965). *The Scientific Analysis of Personality*. Harmondsworth: Penguin Books.

COLQUHOUN, W. P. & CORCORAN, D. W. J. (1964). The effects of time of day and social isolation on the relationship between temperament and performance. *Br. J. soc. clin. Psychol.* **3**, 226–231.

DAVIES, D. R. & HOCKEY, G. R. J. (1966). The effects of noise and doubling the signal frequency on individual differences in visual vigilance performance. *Br. J. Psychol.* **57**, 381–389.

DAVIES, D. R., HOCKEY, G. R. J. & TAYLOR, A. (1969). Varied auditory stimulation, temperament differences and vigilance performance. *Br. J. Psychol.* **60**, 453–457.

EYSENCK, H. J. (1955). A dynamic theory of anxiety and hysteria. *J. ment. Sci.* **101**, 28–51.

EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge & Kegan Paul.

EYSENCK, H. J. (1960). *Experiments with Drugs*. London: Pergamon Press.

EYSENCK, H. J. (1966). Personality and experimental psychology. *Bull. Br. psychol. Soc.* **19**, no. 62, 1–28.

EYSENCK, H. J. (1967). *The Biological Basis of Personality*. Springfield, Ill.: Thomas.

EYSENCK, H. J. & COOKSON, D. (1969). Personality in primary school children. II. Teachers' ratings. *Br. J. educ. Psychol.* **39**, 123–130.

EYSENCK, S. B. G. (1965*a*). *Manual of the Junior Eysenck Personality Inventory*. London: University of London Press.

EYSENCK, S. B. G. (1965*b*). A new scale for personality measurement in children. *Br. J. educ. Psychol.* **35**, 362–367.

FARLEY, F. & FARLEY, S. V. (1967). Extraversion and stimulus-seeking motivation. *J. consult. Psychol.* **31**, 215–216.

FRANCIS, R. D. (1969). Introversion and isolation tolerance. *Percept. mot. Skills* **28**, 534.

GALE, A. (1969). 'Stimulus hunger': individual differences in operant strategy in a button-pressing task. *Behav. Res. Ther.* **7**, 265–274.

GALE, A., COLES, M. & BLAYDON, J. (1969). Extraversion–introversion and the EEG. *Br. J. Psychol.* **60**, 209–223.

HASLAM, D. R. (1967). Individual differences in pain threshold and level of arousal. *Br. J. Psychol.* **58**, 139–142.

HOWARTH, E. (1964). Differences between extraverts and introverts on a button-pressing task. *Psychol. Rep.* **14**, 949–950.

PETRIE, A., COLLINS, W. & SOLOMON, P. (1960). The tolerance for pain and for sensory deprivation. *Am. J. Psychol.* **73**, 80–90.

REED, G. F. & FRANCIS, T. R. (1962). Drive, personality, and audiometric response consistency. *Percept. mot. Skills* **15**, 681–682.

SAVAGE, R. D. (1964). Electro-cerebral activity, extraversion and neuroticism. *Br. J. Psychiat.* **110**, 98–100.

SMITH, S. L. (1968). Extraversion and sensory threshold. *Psychophysiol.* **5**, 293–299.

THOMAS, A. *et al.* (1963). *Behavioral Individuality in Early Childhood*. New York: New York University Press.

WARBURTON, F. W. (1962). The measurement of personality. I. *Educ. Res.* **4**, 2–17.

WEISEN, A. (1965). Differential reinforcing effects of onset and offset of stimulation on the operant behavior of normals, neurotics and psychopaths. (Unpublished Ph.D. thesis, University of Florida.)

(*Manuscript received* 10 *February* 1970; *revised manuscript received* 17 *June* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 381–394
Printed in Great Britain

THE PINMEN RECALL TEST AS A MEASURE OF THE INDIVIDUAL

By P. R. F. CLARKE
Department of Psychiatry, University of Sheffield

The Pinmen Recall Test, an instrument developed as an experimental demonstration of repression, is examined for its applicability as a measure of the individual. Questions are raised about the factors which determine the order in which a subject recalls a set of items which have been seen recently. The first experiment involved eight subjects in an eight-fold [illegible] design. The second was a single test–retest study using [illegible] subjects, all retested at different time intervals. Both studies used balanced orders of presentation and both showed that order of presentation affects order of recall. In both, individual consistency was poor, although group patterns emerged strongly. These group patterns were similar to those found in earlier studies of women students. The material of the second study was examined in the light of a psychiatrist's opinion about what would disturb the individual subject: order of recall did not appear to be related to that opinion.

Forrest & Lee (1962) made a study of repression and perceptual defence using drawings of pinmen or stick figures engaged in activities representative of 10 Murray (1938) needs. In addition to a recall test and tachistoscopic threshold measures with the pictures, their subjects (women students) completed both questionnaire and projective measures of the same 10 needs. A need was deemed 'repressed' when it appeared strongly in the projective test protocol but only weakly in the questionnaire measures. It was predicted that pictures representing such a need would be difficult to see in the tachistoscopic test and relatively late in being recalled in the memory test. Two sets of pictures were prepared (set A and set B) so that the subjects were given different pictures in the perceptual defence and repression parts of the study. Examples of the pictures used appear in Fig. 1. Four groups of subjects were used in a cross-over design with pilot and main replications. The prediction that 'repressed need' pictures would be recalled late was confirmed, but there was no evidence that such needs were defended against in the perceptual experiment. The present paper is concerned only with the recall task.

Clarke *et al.* (1965) gave this recall test to a mixed group of 42 psychiatric patients to assess its viability as a clinical procedure in the hope that it would yield useful information about individual patients. They found that patients differing widely in intelligence, education, socio-economic status and mental health were able to carry out the task. The similarity in the need recall pattern of their patients and Forrest & Lee's (1962) undergraduates was particularly striking. In order to check that the test was really as sensitive to individual differences as the earlier study had suggested, they gave the test again to seven patients who were still available for testing. The test–retest intervals ranged from 16 to 27 days. The same test material was used (some needs from each set) and was presented in random order, i.e. the order of presentation was probably different on retest. Immediately following this retest, the alternative set of material was administered. Of the seven patients only one showed a significant positive test–retest correlation using the same material. Using the alternative material, the only significant test–retest correlation (occasions 1 and 3) was negative.

Although one might expect contamination on repeating a memory test, it would seem reasonable that material which was supposed to be delayed due to repression should again be subject to the same memory block, particularly since it had been learned only to the criterion of one correct recall. Such striking lack of individual consistency gave rise to the present studies of the test as a measure of the individual.

Fig. 1. Examples of Forrest & Lee (1962) drawings. Top row, left to right: set A: Aggression, Harmavoidance, Rejection. Bottom row, left to right: set B: Sex, Exhibition, Rejection. (Reproduced by permission of Forrest & Lee and the American Psychological Association.)

THE PINMEN RECALL TEST

The test is fully described and illustrated in Forrest & Lee's (1962) monograph. The material consists of a pack of 24 cards: 10 of these cards bear drawings of 'stick figures' or 'pinmen' engaged in activities representative of Achievement, Affiliation, Aggression, Dominance, Exhibition, Harmavoidance, Narcism, Nurturance, Rejection and Sex (Murray, 1938), one drawing for each need. The other 14 cards bear drawings of common objects to serve as buffer or dummy items to increase the difficulty of the recall task: they are not considered in scoring. There are two parallel sets of the need pictures but the same dummies are used for both sets. Ordinarily the need cards and dummy cards are randomized beforehand in the pack with the exception that the pack always starts and ends with at least two dummies.

The subject's task is to look through the pack of 24 cards one at a time and then to describe as many of them as he can recall. Except for end dummies, all cards described are removed. Remaining items are presented again in the same manner and the test continued until all 24 cards have been recalled. In effect, this is learning the set of need

cards to the criterion of one correct response per need. No time constraints or controls are imposed: the subject is asked to look at each card 'to take in its meaning'. Since he holds the pack of cards and leafs through them one at a time he rarely spends more than a few seconds looking at each card. As soon as he has finished looking through the pack, it is removed from sight and he is immediately asked to describe as many of the cards as he can remember in any order. When he can recall no more, the pack is prepared for the next trial by removing all recalled cards (except the end buffers) and then presented again in the same order. The present procedure is slightly different from that of Forrest & Lee (1962) (who re-randomized the cards between trials) to allow specific examination of order of presentation effects.

Accepting the repression interpretation, the rank order in which the 10 need cards are recalled (disregarding the dummy items and the number of exposures used) can be taken as a measure of the relative accessibility to the subject of these 10 needs. Even if the repression interpretation were not acceptable, it would be possible, although not necessary, to take the view that such a rank order affords a description of the individual's need structure. The rank correlation between such orders obtained from a single subject on different occasions yields a measure of individual consistency, while rank correlations (or coefficients of concordance) between the orders of two or more subjects measure similarities between individuals. The consensus rank orders for different groups of subjects can of course also be compared by rank correlation measures.

Before describing the experiments in detail, consideration should be given to the sources of variation which are likely to influence the order in which an individual recalls a set of such materials and which might account for the high inter-subject and low intra-subject consistency previously found. Some of these factors may easily be balanced out in an experimental group but be very difficult to control in an individual case. The possibility that the order of recall on this test is purely random can be excluded from the outset: since group patterns emerged for the need cards they are clearly not recalled haphazardly; this is supported by the further finding that the dummy cards failed to show significant patterns of recall over the groups (Forrest & Lee, 1962). The test measures something.

The fact that groups of subjects showed significant patterns in their recall orders for the need cards (Forrest & Lee, 1962) implies that the cards, or the activities depicted on the cards, were equally memorable or 'reportable' to many subjects. This might be because the cards differ in their complexity; in the complexity of the drawing, in the difficulty of finding words to describe the activity, in the difficulty of grasping the nature of the interaction between the pin-figures. The activities depicted might be differentially impolite to report, or differ in their interest or value, or their frequency of occurrence. The needs intended to be represented might be differentially urgent or postponable, active or passive, satisfiable in few or many ways, frequently or rarely displayed, culturally desirable or taboo within the value system of a particular culture. Such factors could conceivably account for agreements between many subjects: groups of subjects having clearly different cultural or subcultural backgrounds might be expected to show patterns differing in proportion to their cultural differences.

Subjects would also differ from each other in many ways. However, as long as such differences were stable characteristics of the individuals—different interests, values,

needs, skills in perceiving and reporting—there should be high intra-subject consistency in the recall of the pictures. Such individual patterns would of course only be discernible as individual in so far as they deviated from the general group or cultural patterns. This raises similar issues as the disentanglement of lack of pathology from social desirability responding in questionnaires like the MMPI.

So far only factors producing consistency have been considered. Several influences appear to work to produce inconsistent patterns when the test is administered more than once to the same subject. For example, if the experimenter presents the material in different sequences on two occasions this imposes an organization on the material independent of the subject's characteristics and might account for low reliability. Lee (personal communication) was not able to find any evidence that order of presentation had influenced order of recall, but his technique of randomizing the pack afresh for each subject confounded order of presentation and individual differences, so that slight effects might have been imperceptible. The present studies explicitly examined this effect.

Again, if the order of recall is a measure of the individual's value system, any change in that value system should be reflected in a different order of recall on subsequent testing, but major changes in a value system are not likely in the short run, particularly changes in long-established codes of censorship. On the other hand, quite a short time interval might produce changes in the urgency hierarchy of a subject's needs rather than his hierarchy of their objectionable quality. Primary needs certainly compete for direction of the individual's activity, depending on how urgently they need to be met. Even extreme randiness is likely to yield to extreme hunger, and nothing much is likely to matter more than the need for air if we are suffocating. Similarly it might be expected that the varying frustrations and satisfactions of day-to-day living might be reflected in changes in the strength of needs like affiliation or dominance or exhibition.

Other sources of inconsistency lie in the fact that the material may change for the subject once he has done the test. Although the items are all learned to the criterion of one correct response, they will have been exposed a varying number of times so that they may be differentially accessible. Items delayed the first time may acquire salience simply through their delay: these two factors may interact with evaluative factors to produce very complex carry-over effects which might nevertheless remain lawful (Lee, personal communication). Moreover, repeated exposure of the material might desensitize the subject to its objectionable nature. Further sources of inconsistency are of the error of measurement sort. Thus small samples of recalling are less likely to be stable than large samples, and a measure like serial order of recall may show big score differences due to quite small changes in memorability because the responses must queue for emission through a single output channel.

There is already some evidence bearing on these points. In the original study Forrest & Lee (1962) found no significant differences in either the tachistoscopic or recall measures between the two forms of the test. Moreover, although the pictures differed among themselves for accuracy of tachistoscopic perception and for speed of recall, there were no systematic similarities between ease of seeing and recalling the pictures. Thus mere complexity or obscurity of the pictures is unlikely to account for the overall order of recall which has emerged from the studies so far. As a further

check on this point the author asked a group of female hospital staff to rank the pictures (within each set) for 'ease of description'. Although his subjects agreed highly significantly with each other and showed a clear hierarchy of ease of description, their consensus rankings for the two sets were not significantly related to the Forrest & Lee (1962) orders of recall. Moreover, the consensus orderings of the two sets were not significantly correlated ($\rho_s = 0{\cdot}44$).

Inspection of the pictures makes the notion that the recall pattern was due to social taboos on reporting implausible. The only picture remotely likely to be given an impolite description is the Sex picture in set A, and this has almost invariably been well-remembered. On the other hand, the main Forrest & Lee (1962) finding that delayed recall was related to the projective and questionnaire measures supports the view that the pictures have a hierarchy of emotional acceptability. This probably transcends day-to-day vicissitudes of drive strengths, since the personality measures were completed in the subjects' own time after the recall task. Although the case that the test is sensitive to repression is elegantly argued by Forrest & Lee (1962), this is not the only possible value it might have. Any consistent relationship between a subject's value system and such a brief test would be of interest.

The first experiment in the present study assumed that the subject's stable value system is the most important factor in the test and sought to eliminate some possible sources of individual inconsistency. It did this by testing for possible order of presentation effects and by repeating the test to bring out any individual pattern consistent enough to stand out from general cultural values. The second experiment then had to take more seriously the carry-over, detoxication and personal change views and attempt to choose between them.

Experiment I

Method

Eight in-patients and day-patients of the psychiatric unit of a teaching hospital were selected as being as unlike women undergraduates as possible to allow subcultural effects to be examined. All agreed to carry out the tests. All were married men over the age of 38, of elementary education and low occupational status. Only one had a Mill Hill Vocabulary score above the mean for his

Table 1. *Orders of presentation based on Forrest & Lee* (1962) *orders of recall*

	Needs									
	Ach	Aff	Agg	Dom	Exh	Harmav	Nar	Nur	Rej	Sex
Set A orders										
A	7	6	1	5	10	4	8	3	9	2
B	4	5	10	6	1	7	3	8	2	9
C	9	10	1	5	6	4	8	3	7	2
D	4	5	6	10	1	9	3	8	2	7
Set B orders										
E	7	3	2	8	10	1	6	5	9	4
F	4	8	9	3	1	10	5	6	2	7
G	4	8	7	3	1	6	5	10	2	9
H	9	3	2	8	6	1	10	5	7	4
Mean rank	6	6	4·75	6	4·5	5·25	6	6	5	5·5

Kendall's coefficient of concordance (W) = $0{\cdot}04$, $P > 0{\cdot}90$.
Mean inter-order $\rho_s = -0{\cdot}10$.

age (Raven, 1948); only four had WAIS Block Design scores above the mean for their age (Wechsler, 1955). In addition to their diagnosis of major psychiatric illness, six were noted to have abnormal personality.

In this study, the test was given eight times to each subject, using set A and set B pictures alternately. The eight predetermined orders of presentation (Table 1) were based on systematic variations of the orders of recall observed in Forrest & Lee's (1962) main experiment. The orders of presentation were designed to give a wide range of intercorrelations while balancing out within each subject. Each subject received the eight orders of presentation in a different sequence. Except for the first subject (tested over 11 days), all eight tests were given within 5 days, usually twice a day with one day's rest. The intervals between tests on the same day varied from half an hour to several hours and were usually filled with other psychological testing for descriptive or clinical purposes.

If order of presentation has any important effect, there should be consistency within orders across subjects. If the test is sensitive primarily to individual patterns, there should be strong consistency within individual subjects and little agreement between subjects. If it is sensitive mainly to subcultural patterns, there should be considerable agreement between subjects but little agreement between their overall pattern and the Forrest & Lee (1962) patterns. If it is primarily sensitive to the overall pattern of the culture, these subjects should agree with each other and also with the earlier findings.

Results

Order of recall is affected by order of presentation in complex ways, as shown by Table 2. The second column gives values of Kendall's coefficient of concordance for the inter-individual agreement in recall order invoked by each of the eight orders of

Table 2. *The effect of order of presentation on order of recall*

Order of presentation	Recall concordance (Kendall's W)	Order of presentation × recall consensus		Order of presentation × Forrest–Lee recall	
		ρ_s	Rank	ρ_s	Rank
Set A					
A	0·22	+0·63*	8	+0·99**	8
B	0·20	−0·59	2	−0·99**	1
C	0·26*	+0·14	7	+0·77**	7
D	0·09	−0·32	3	−0·75**	3
Set B					
E	0·27*	+0·01	5	+0·52	5
F	0·33**	−0·64*	1	−0·52	4
G	0·29*	−0·21	4	−0·76**	2
H	0·28*	+0·13	6	+0·76**	6

* $P < 0·05$. ** $P < 0·01$.

presentation. Those values showing significant agreement imply that order of presentation does affect order of recall. However, the various orders of presentation are not all related in the same way to the consensus of orders of recall which they invoked, as shown by the correlations in column 3: there are both significant and insignificant, positive and negative relationships. Column 5 gives the values of ρ_s holding between each order of presentation used and the overall recall order found by Forrest & Lee (1962). Inspection of the size and signs of the correlations in columns 3 and 5 and the fact that the ranks in columns 4 and 6 are highly significantly related ($\rho_s = +0·89$) demonstrate that the power a particular order of presentation has to determine its order of recall is a function of the similarity it bears to the original orders of recall

found by Forrest & Lee (1962), again emphasizing the non-randomness of the original recall orders.

The data in Table 3 demonstrate a tendency towards self-consistency for all subjects, but the effect is not very marked and reaches a significant level in only half the subjects. These differences in consistency were not obviously related to other characteristics of the individuals.

A measure of the strength of order of presentation effect relative to individual patterning can be obtained by comparing the eight values for Kendall's coefficient of concordance in Table 2 with those in Table 3 using the Mann–Whitney test. This gives $U = 23$, $P < 0{\cdot}38$, showing that these two sources of variation do not differ in strength.

Table 3 *Individual consistency of recall over eight occasions (Kendall's coefficient of concordance)*

	Subjects								
	1	2	3	4	5	6	7	8	Overall
W	0·31**	0·25*	0·27*	0·22	0·19	0·15	0·27*	0·17	0·10***
Mean ρ_s between trials	0·21	0·14	0·16	0·11	0·07	0·03	0·16	0·05	0·10

* $P < 0{\cdot}05$. ** $P < 0{\cdot}01$. *** $P < 0{\cdot}001$.

In contrast to these findings, the concordance over all 64 recall orders (eight orders for eight subjects) showed very significant agreement ($W = 0{\cdot}10$, $P < 0{\cdot}001$). Nevertheless the average rank correlation between each of these orders was extremely small (ρ_s av. $= +0{\cdot}10$). Correlations between the present consensus and earlier studies show that this agreement is not specific to this subcultural grouping. Indeed subcultural similarities do not appear to be accurately reflected in the similarities between recall orders for the various groups studied, as shown by Tables 6 and 7 where successive samples of students and of psychiatric patients seem to resemble the other population rather than their own.

Thus it appears that all three factors (order of presentation, individual patterns and cultural patterns) play a part in determining the order in which a subject recalls the pictures on any occasion. None of these sources of variation, however, appear very strong and it would seem either that other factors are involved as well or that there is a very large error component.

Differences between the two sets of pictures might account for some of this variation. Indeed set B orders of presentation yield higher coefficients of concordance than set A (Table 2, column 2, Mann–Whitney $U = 0$; $P < 0{\cdot}03$), and the consensus of the 32 set A occasions (four orders for each of eight subjects) is not significantly correlated with the consensus of the 32 set B occasions ($\rho_s = 0{\cdot}41$). Subjects also showed a trend to greater self-consistency when recalling their set B material. These findings are not in keeping with the Forrest & Lee (1962) study, which revealed no important differences between sets, but they agree with the author's study of the ease of description of the cards.

The presence of carry-over effects due to repeated testing may have obscured other effects. Some evidence for this can be found by looking at the number of trials needed

to meet the learning criterion on successive occasions. There was a slight steady improvement for the group as a whole over the first five occasions, but the last three occasions needed extra trials and only two of the eight subjects took fewer trials on the last than on the first occasion.

This experiment has clearly shown that order of presentation is a variable which must be controlled before an individual pattern can be reliably derived from the test, but it has not provided such a reliable measure. Whether this failure is due to inadequate balancing of order of presentation, to the carry-over or desensitization effects of repeated testing, to stronger cultural patterning or to sheer instability of patterns of recall requires further study.

The fact that these subjects showed so little 'savings' supports carry-over rather than desensitization to any objectionable emotional tone of the pictures since it seems likely that repeated learning plus desensitization would have made the learning task easier than in fact it became. There is a four-horned dilemma here: single retesting will be contaminated by order of presentation; multiple retesting offers an opportunity for desensitization and carry-over; short-term retesting maximizes carry-over effects; long-term retesting allows time for the individual's personal values to change.

Experiment II*a*

The purpose of this further experiment was to test the hypothesis that carry-over effects are responsible for the low individual consistency previously observed.

Method

The main feature of the design lies in the retesting of each subject at different intervals of time, since carry-over effects can be expected to decay with the passage of time. Desensitization was minimized by single retesting. Individual consistency was encouraged by using the more self-consistent set of pictures (set B) and by presenting the cards in the same order both times. A weakness of the study is that changes in personal values can be expected to increase with the passage of time; however, time should have opposite effects on carry-over and personal change factors.

Twenty in-patients and day-patients of the same teaching hospital were selected as being likely to remain available for the appropriate time lapse, as likely to cooperate, as not likely to be given ECT, and as showing no gross evidence of intellectual impairment. Three men and 17 women completed the testing. Their ages ranged from 17 to 63 with a fairly even spread about a median of 39 years. They showed a wide range of educational and occupational achievement and bore a variety of psychiatric diagnoses.

Each subject took the Pinmen Recall test twice: for 10 subjects the retest intervals (all different) ranged from 1 to 10 days and for the other 10 subjects from 21 to 30 days. The orders of presentation were again based on systematic permutations of the order of recall of set B pictures found in Forrest & Lee's (1962) main experiment, each picture appearing twice in each position. Each subject had his own order of presentation which was used for both test and retest. Over the 20 subjects the orders of presentation balanced out to show no concordance.

Results

Spearman rank correlations were calculated for each subject comparing his initial and retest recall orders. As column 2 in Table 4 shows, only two of these reached a significant level and their average value is very low and of dubious significance. By the criteria of Sakoda *et al.* (1954) the present data may have occurred by chance, but Taylor & Fong's (1953) method makes this average value significant at $P < 0{\cdot}01$.

Correlating the data in columns 2 and 3 of Table 4, we find that the stability of the individual's pattern on test and retest is not a function of the time lapse involved ($\rho_s = -0{\cdot}006$). Further, the test–retest correlations are not different for the 10 short and 10 long intervals (Mann–Whitney $U = 42$). This runs counter to the hypothesis that earlier inconsistency was due to carry-over effects (consistency should improve with time) and also counter to the hypothesis that it was due to changes in personal values (consistency should decrease with time).

Table 4. *Individual consistency*

S	Test–retest correlations ρ_s	Intervals (days)	Savings (trials)	Order of presentation × Forrest–Lee recall ρ_s
1	+0·27	1	0	+0·28
2	−0·47	2	1	−0·28
3	+0·01	3	2	−0·02
4	+0·39	4	2	+0·45
5	+0·37	4	2	+0·03
6	+0·38	6	1	−0·08
7	+0·27	7	0	−0·44
8	+0·08	8	1	−0·99**
9	−0·14	9	4	−0·46
10	−0·08	10	1	−0·28
11	−0·07	21	1	−0·45
12	+0·58*	22	1	+0·46
13	+0·19	23	0	−0·46
14	−0·04	24	0	−0·52
15	+0·84**	25	1	+0·99**
16	+0·49	26	1	+0·43
17	+0·06	27	0	+0·28
18	+0·16	28	0	+0·46
19	−0·05	29	1	+0·09
20	+0·09	30	1	+0·53
Mean	+0·17**			

* $P < 0{\cdot}05$. ** $P < 0{\cdot}01$. *** $P < 0{\cdot}001$.

Another test of the carry-over hypothesis can be made by comparing the amount of savings in reaching the learning criterion on retest shown by the 10 most and 10 least consistent subjects (column 4, Table 4), since the less the subject can remember of the first test, the less it should interfere with the second. Not surprisingly there was some savings in the group as a whole, but the Kolmogorov–Smirnov test showed no significant difference between the consistent and inconsistent individuals. However, the intervals may have been too brief to constitute an adequate test of the carry-over hypothesis, since no significant differences in savings emerged between the short and long intervals either.

Order of presentation is again a factor to be reckoned with in that test–retest stability is higher the more closely the order of presentation resembles Forrest & Lee's (1962) order of recall (set B, main experiment) ($\rho_s = +0{\cdot}57$; $P < 0{\cdot}01$). This would suggest that individual differences in recall orders are basically unstable and negligible in comparison with cultural patterns and the order of presentation.

However, there appear to be more complex but nonetheless lawful relationships between the ease of remembering a need on the first occasion and its recall position on

the second. Certain recall positions on initial testing were consistently associated with other recall positions on retest (Friedman two-way analysis of variance, $P < 0\cdot05$). The pattern found is shown by the consensus retest recall ranks given in Table 5: this shows that items recalled both first and last on occasion one tend to be late on retest. This may represent an interaction between number of times a card is shown and a basic set to delay or facilitate its recall, as suggested by Lee (personal communication). There is, however, little evidence that the number of exposures needed to recall a card the first time is reflected in its retest recall rank. Comparisons of the frequency distributions of retest recall ranks for cards needing one, two, three or more exposures on the initial test yielded only one significant difference: cards exposed once were recalled earlier on retest than cards needing two exposures (Kolmogorov–Smirnov, $D = 0\cdot30$; $P < 0\cdot01$).

Table 5. *Retest recall as a function of test recall rank*

Initial position	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Sum of retest recall ranks for 20 subjects	128	92	96	84	85	124	105	121	133	132
Consensus retest recall rank	8	3	4	1	2	7	5	6	10	9

Friedman $\chi^2_r = 19\cdot1$, d.f. $= 9$, $P < 0\cdot05$.

Yet another attack on this question of 'differential exposure' as a factor in change of position can be made by tabulating changes towards later or earlier position on retest for all the items recalled on a particular trial and comparing these figures against 'expected' values (counting the possibilities for a card in a particular position to occur earlier or later: the card recalled first can only occur in the same place or in one of nine later places on retest; the second card could occur in one earlier position or eight later positions). The observed minus expected values were compared separately for 'no change', 'change to earlier' and 'change to later' for subjects batched by number of trials they had needed on first testing. Only one of the nine calculations reached $P < 0\cdot05$: the subjects who took four trials showed significantly less retest stability ('no change') than expected on the cards they had recalled on the first test trial (Friedman two-way analysis of variance).

It should also be remarked that those subjects who differ most from the consensus recall order on the first occasion tend to remain different from the second consensus on retest, while subjects closely related to the consensus the first time remain close the second time ($\rho_s = +0\cdot42$; $P < 0\cdot05$, using as scores the sums of d^2 between individual and consensus ranks on the two occasions).

Thus while the evidence suggests that there may be consistent relationships between the first and second order of recall despite the low value of ρ_s for most individuals, the cause of such changes remains obscure.

As in previous studies, despite the individual changes, the subjects agreed with each other on both occasions (Kendall's coefficient of concordance yielding $P < 0\cdot001$ and $P < 0\cdot05$ respectively) and the test–retest correlation of their consensus orders of recall reached $\rho_s = +0\cdot72$, $P < 0\cdot05$. Relationships with earlier consensus orders are again strong and puzzling as shown by the data in Tables 6 and 7. The four Forrest &

Lee (1962) samples appear more like some of the psychiatric samples than they do like each other and groups seeing different sets are sometimes more alike than groups seeing the same sets.

Table 6. *Orders of recall (consensus ranks)*

Sample	Needs: Ach	Aff	Agg	Dom	Exh	Harmav	Nar	Nur	Rej	Sex
Forrest & Lee										
Pilot										
Set A	3	4	8	9	6	7	5	2	10	1
Set B	5·5	3	9	4	8	10	7	1	5·5	2
Main										
Set A	7·5	6	1	5	10	4	7·5	3	9	2
Set B	8·5	3	4	8·5	6	5	10	1	7	2
Clarke *et al.*										
Set AB	7	1·5	3	6	8	10	9	4	5	1·5
Expt. I										
Set A	7	2	1	9	5	4	8	6	10	3
Set B	4	3	8	2	9	1	5	6	10	7
Expt. II										
Test, Set B	5	2	6	4	10	9	7	3	8	1
Retest, Set B	7	3	6	1	9	8	4	2	10	5

Table 7. *Correlations between consensus recall orders of several samples* (ρ_s)

	Forrest & Lee					Present			
	Pilot	Main			Clarke *et al.*	Expt. I			Expt. II
	Set B	Set A	Set B	Sets A+B	Set AB	Set A	Set B	Sets A+B	Set B
Forrest & Lee									
Pilot									
Set A	+0·56	+0·25	+0·47	.	.	+0·35	.	.	.
Set B	.	+0·19	+0·43	.	.	.	+0·01	.	+0·84**
Sets A+B	.	.	.	+0·48	+0·65*	.	.	+0·92**	.
Main									
Set A	.	.	+0·62*	.	.	+0·58*	.	.	.
Set B	.	.	.	.	.	.	−0·14	.	+0·49
Sets A+B	.	.	.	.	+0·72*	.	.	+0·61*	.
Clarke *et al.*									
Set AB	.	.	.	.	.	.	.	+0·22	.
Present Expt. I									
Set A	.	.	.	.	.	.	−0·41	.	.
Set B	.	.	.	.	.	.	.	.	+0·22

* $P < 0{\cdot}05$. ** $P < 0{\cdot}01$.

Experiment II*b*

The data from the initial testing in Expt. II*a* were also used for a cursory examination of the relationship between the psychopathology of the individual subject and his recall order.

Method

After the collection of the test data a psychiatric registrar who personally knew all the patients examined their case-notes. For each subject, he was asked to select two of the pictures as 'most

likely to disturb the patient at a conscious level' and two 'most likely to be associated with the [illegible] up forbidden instinctual demands'. He was allowed to nominate the same picture for each category if he wished; in fact, for eight subjects there was no overlap between categories, and for the other 11 subjects one need each was nominated as forbidden but not consciously disturbing. The recall rate of these needs thus nominated to be emotionally loaded was examined in three ways: by counting the number of exposures needed to recall these needs, by comparing the distribution of the observed frequencies of recall ranks for these needs against various hypothetical distributions, and by comparing their recall ranks with the group consensus for individual deviance.

Although these particular labels might not clearly reflect pure hypotheses about repression they were chosen as in keeping with the realities of psychiatric practice. Evidence for competing hypotheses that such material might be delayed in recall ('repression'), or facilitated in recall, or both, was examined.

Results

Firstly, no association could be demonstrated between the psychiatrist's labelling of a picture and the number of exposures required for it to meet the learning criterion (χ^2 tests). Generally, however, the needs more frequently nominated by the psychiatrist were those less often requiring a second exposure.

Secondly, the Kolmogorov–Smirnov test was used to compare the observed frequency distributions of the 40 needs in each of the psychiatrist's categories over the 10 recall ranks against various hypothetical distributions. Six hypothetical distributions of the 40 needs were used: one rectangular distribution (four items in each of 10 positions), two 'repressing' distributions (bunching the 40 equally into the last five positions or the last two positions), two 'alerting' distributions (bunching the 40 equally into the first two positions or the first five positions) and one 'repressing and alerting' distribution (bunching them equally into the first two and the last two positions). The 40 'consciously disturbing' needs differed significantly ($P < 0{\cdot}01$) from the last five hypothetical distributions but matched the rectangular distribution quite closely. The 40 'forbidden' needs not only differed at the $P < 0{\cdot}01$ level from the last five hypothetical distributions but even deviated at the $P < 0{\cdot}10$ level from the rectangular distribution. Calculations based on only those 'forbidden' needs which were not 'consciously disturbing' again showed the best fit with the rectangular distributions.

Thirdly, the nominated needs were examined for their initial deviance from the group recall pattern at the suggestion of Lee (personal communication). The individual's pattern of values might show up more clearly if the ranks of the group recall order were subtracted from each of the individual's recall ranks. Thus an individual recalling the Sex picture last might be expected to have strong individual reasons for doing so, since the group as a whole recalled it first. The distribution of the observed deviations from the consensus for the nominated needs was compared with an expected distribution calculated from the number of ways differences of each size from $+9$ to -9 could be obtained from two sets of the ranks 1–10. Neither the 'consciously disturbing' nor the 'forbidden' needs differed significantly from the expected distribution (Kolmogorov–Smirnov test) showing that the needs nominated for an individual were not noticeably deviant from the group pattern.

While the opinion of a single psychiatrist with relatively slight knowledge of the patients tested is only a crude index compared with Forrest & Lee's (1962) psychometric criterion of repression in normal subjects, these findings must be held to weaken any claim that the Pinmen Recall test is sensitive to which needs a patient is

likely to repress. It is of interest that the different criteria of what is personally 'objectionable' in the two studies throw up different needs, while the two sets of subjects recall the pictures rather similarly. For example, Harmavoidance was nominated by the psychiatrist in only four of the 20 patients but reached the repression criterion in all batches of Forrest & Lee's (1962) students. Nurturance was nominated for eight of the patients but was 'repressed' in only one batch of students. Harmavoidance was, however, recalled ninth by the patients and eighth by the students, and Nurturance third by the patients and second by the students.

It is not claimed that repression does not occur, nor that a psychiatrist must, by virtue of his office, be able to predict how emotional material will be handled. It is, however, suggested that the Pinmen Recall test is not obviously sensitive to clinically disturbing material.

Conclusions

The present experiments demonstrate that the Pinmen Recall test does not in its present form provide a stable measure of the individual over the course of a few days or weeks and suggest that it is not obviously sensitive to clinically disturbing material despite the evidence that it was sensitive to individual patterns of motivation in the original monograph by Forrest & Lee (1962). The present studies have established that order of presentation is a factor which must be controlled but it cannot entirely account for the intra-subject instability found.

The generality of the Forrest & Lee (1962) orders of recall has been established and their potency to influence the order of presentation effect demonstrated. This generality and potency clearly imply that, despite individual unreliability, the test is measuring something of interest.

The findings that individual stability over time is a function of neither the test–retest interval (up to a month) nor of the savings shown in relearning the material weakens but does not entirely refute the hypothesis that the instability is due to carry-over effects. An alternative hypothesis, that it is due to desensitization to the objectionable material, is not convincing since the material in the second experiment had only been learnt once to a criterion of one correct description per picture. A further hypothesis is that the subject is changing his scale of values during the interval. Since these test–retest studies have been carried out on patients under treatment this has some attraction. However, it would be optimistic to suppose that any major rebuilding of personality patterns would have occurred over the relatively short intervals studied. Moreover, if this were the case there should be a negative relationship between stability and length of test–retest interval (the greater stability over the shorter interval), but there is not unless it has been balanced by carry-over factors.

Since the test is short and the hierarchy depends simply on the order in which the pictures are described from memory but once, it may simply be that the present scores derived from the test are unreliable measures of stable, and important, individual characteristics. It would therefore be useful to devise a longer test using more needs, more pictures for each need, a bigger sample of remembering behaviour and better control of order of presentation. Such a test might also attempt to reduce the ambiguity of the present pictures which allow the subject to identify with either the active or passive person involved (e.g. Dominance: 'a man being arrested' or 'a

policeman arresting a man'). A new larger set of pictures has been devised and tested with additional checks on the stability of the subject's value system and this will be reported later. The possibility that it is not the subject's value system which changes but rather the moment-to-moment urgency of satisfaction of the needs could then be examined with an improved instrument manipulating the strength of specific needs experimentally.

The author wishes to express his thanks to Professors Forrest and Lee for the test material and for many useful suggestions, Dr M. [illegible] for the psychiatric opinions in Expt. II, the Medical Research Council for a grant (G 964/281/C) for computational assistance, Mrs D. [illegible] and Mrs D. [illegible] for carrying out most of the calculations, and colleagues in the Department of Psychology, University of Sheffield, for helpful criticism. The illustration in Fig. 1 is reproduced from the original monograph by kind permission of Professors Forrest and Lee and the American Psychological Association.

REFERENCES

CLARKE, P. R. F., HOWE, M. J. A. & LOXLEY, F. D. (1965). Some reflections on the Forrest Recall Test. *Bull. Br. psychol. Soc.* **18**, no. 59, 17A.

FORREST, D. W. & LEE, S. G. (1962). Mechanisms of defense and readiness in perception and recall. *Psychol. Monogr.* **76**, 4. (Whole no. 523.)

MURRAY, H. A. (1938). *Explorations in Personality*. New York: Oxford University Press.

RAVEN, J. C. (1948). *Guide to Using the Mill Hill Vocabulary Scale*. London: Lewis.

SAKODA, J. M., COHEN, B. H. & BEALL, G. (1954). Test of significance for a series of statistical tests. *Psychol. Bull.* **51**, 172–175.

TAYLOR, W. L. & FONG, C. (1963). Some contributions to average rank correlation methods and to the distribution of the average rank correlation coefficient. *J. Am. statist. Ass.* **58**, 756–769.

WECHSLER, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.

(*Manuscript received* 13 *August* 1968; *revised manuscript received* 19 *June* 1970)

Br. J. Psychol. (1971), 62, 3, pp. 395–401
Printed in Great Britain

DETECTABILITY OF ITEMS IN THE EYSENCK PERSONALITY INVENTORY

By R. P. POWER and K. D. MacRAE

Department of Psychology and Department of Medical Statistics,
The Queen's University, Belfast

[illegible] students completed Form B of the Eysenck Personality Inventory, and were then [illegible] which items were measures of neuroticism (N), extraversion (E) and lying (L). [illegible] were correctly detected, and detectability of items as N and E items was [illegible] with their factor loadings. Correct detection of N items was not related to N factor [illegible] but correct detection of E items was related to E factor loadings. This latter finding [illegible] to be due to the unsatisfactory nature of the E scale. Ability to detect items was not related to personality or intelligence. It was concluded that a major difficulty exists, [illegible] in terms of its factor loading the more detectable it is.

In recent years it has been shown that subjects are capable of simulating any of a variety of personality characteristics when completing personality inventories, and that they are, to some extent, able to avoid detection by the lie scales often included in such inventories (Power, 1968; Power & O'Donovan, 1969). This ability is not restricted to inventories, and can be applied when physiological measures which relate to personality are taken (Power & Thompson, 1970).

So far little attention has been paid to establishing what characteristics of the items or of the subjects allow the detection of items as measures of the relevant personality variables or as lie scale items. The investigation reported here was designed to do just this. Subjects were given copies of the Eysenck Personality Inventory (EPI) (Eysenck & Eysenck, 1964; Eysenck *et al.*, 1969), told what the three scales were, and asked to indicate the scale to which the item belonged. The relative detectability of the items was then correlated with their factor loadings, and the ability of the subjects to detect the items was correlated with measures of personality and intelligence.

Method

Procedure. The subjects were asked to complete Form B of the EPI in the usual fashion. After they had done this they were given a fresh copy and asked to write on the back just what they thought it was measuring. They were then given the following instructions:

'Each item in the EPI measures one of three things: neuroticism (N), extraversion (E) or lying (L). Lying is intended here to be pretending to be a very stable, non-neurotic person indeed. You are asked to read each item in the test and decide what personality trait the question is supposed to measure, and whether a "Yes" or "No" answer indicates that personality trait. Record your decision by putting N, E or L in either the "Yes" or the "No" circle after the question. Thus if you decide that saying "No" to the question "I enjoy travelling by train" indicates extraversion you would put E in the "No" circle after that question. If you cannot decide to which scale an item belongs make the best guess you can.'

The inventories given before this were scored for N, E and L in the usual fashion, and scores for the AH 5 (Heim, 1956) administered on another occasion were obtained.

Subjects. The subjects were 55 men and women students in an introductory course in psychology, and the testing was done during their first laboratory class. From these 50 usable EPIs were obtained: the remaining five did not attempt all items and were discarded from the subsequent analysis.

Design. There were to be several stages in the analysis: (1) to establish whether the subjects, [illegible] the EPI [illegible] (2) to [illegible] whether the [illegible] N, E or L [illegible] chance; (3) [illegible] whether the detection of the items as N, E or L [illegible] (4) [illegible] N, E, and L and to correlate [illegible] on the AH 5; (5) to correlate detectability of the N and E items with their factor loadings.

RESULTS AND DISCUSSION

Detectability of the scales

Of the 50 subjects, 36 correctly guessed that the inventory was designed to measure N, 29 that it was designed to measure E, and four subjects stated that it measured if a person was a liar. It was not clear whether lying was intended to mean that the subjects saw that the inventory contained a lie scale, or whether they meant that truthfulness in a more general sense was being measured. No subject guessed the presence of all three scales, although 20 guessed both N and E. In addition, 10 subjects produced guesses that other measures were being sought by the inventory; for example, prejudices, aggression and sense of humour.

Detectability of the items

Mean frequencies of guesses* are reported in Table 1. The EPI has 57 items, 24 of which when answered Yes contribute a score of 1 to the N total (N Yes items), 15 when answered Yes contribute to the E total (E Yes items), 9 when answered No contribute to the E total (E No items), 3 when answered Yes contribute to the L total (L Yes items) and 6 when answered No contribute to the L total (L No items).

Table 1. *Mean frequencies of guesses of EPI items*

	Guesses					
Items	Yes N	No N	Yes E	No E	Yes L	No L
N Yes	29·9	0·8	5·1	2·9	1·2	10·1
E Yes	0·6	4·5	35·1	0·9	4·5	4·5
E No	13·0	1·4	4·3	24·4	3·3	3·4
L Yes	2·3	1·0	2·0	2·0	38·0	4·7
L No	3·5	0·8	6·2	0·5	1·7	37·3

Each item in each of the above five categories of items was subjected to a χ^2 analysis in order to determine whether the distribution of guesses for each item departed significantly from chance ($f_r = 8·33$ for each guess). All χ^2 values indicated that the distribution of guesses in the case of all 57 items departed from chance ($P < 0·001$ in each case).

Each item was then examined to establish whether it was detected correctly. All N Yes items were detected correctly, χ^2 giving $P < 0·001$ except for items 7, 28 and 52: $P < 0·05$ for item 7, and items 28 and 52 were detected as E items. In addition to being guessed as N Yes items, two more were guessed to be E items, and eight items, correctly guessed as N Yes items, were also guessed as L No items. This rate

* Detailed analyses of the results, item by item, listing frequency of correct detections, frequency of incorrect detections, items which are significantly more or less detectable and the factor loading of the E and N items will be sent on request with reprints.

as guessing N Yes items to be L items when answered No suggests that what is one person's neurotic symptom is a lie if denied by another.

All the E Yes items were correctly detected with $P < 0{\cdot}001$. In addition to being guessed as E Yes items one item was guessed as being an N No item, one as an L Yes item and two as L No items. All the E No items were correctly detected with $P < 0{\cdot}001$ except for items 20 and 30 for which $P < 0{\cdot}01$. Four items were significantly misdetected as being N Yes items.

All the L items were correctly detected with $P < 0{\cdot}001$, and none of these was significantly misidentified.

Clearly, on the basis of these results, it can be concluded that most of the items in the EPI are readily detectable by intelligent and psychologically fairly naive subjects.

Characteristics of the subjects

There were individual differences in ability at detecting the items. It was thought that this might be related to verbal intelligence, or that there might be a relationship between scores on the scales and ability to detect items of that type. Therefore scores on the three scales, ability to detect items of that type, and scores on both parts of the AH 5 were intercorrelated. These correlations showed little of interest (i.e. the high correlation between the two parts of the AH 5 was not relevant). Not surprisingly, the ability to detect one type of item correlated with ability to detect another type, all these correlations being positive, and 2/3 being significant. There were no significant correlations between scores on the three scales and ability to detect items of that type. Surprisingly, there were significant negative correlations, each $-0{\cdot}29$, between Part II (spatial) of the AH 5 and detection of E No and L Yes items. Since these correlations were low, had no face validity and were two of many in a large correlation matrix it was decided to treat them as Type 1 errors. We may be making the wrong decision here, but are unable to account for the correlations in any other way.

The results from this portion of the analysis were somewhat disappointing, although it is reasonable to suppose that a correlation between verbal ability and ability to detect items exists when the range of intelligence is not so restricted.

Characteristics of the items

Before a more detailed analysis of the items scale by scale, correlations between guesses of the nature of the 48 items of the N and E scales and their factor loadings (Eysenck *et al.*, 1969) are considered. They are shown in Table 2. In this case no account is taken of the correctness of the detection of an item as a certain type.

The N factor loadings of the items are highly correlated with guesses that the item is an N Yes item and negatively correlated with guesses that the item is an N No item. Thus the higher the factor loading on N the more likely it is that the item will be identified as an N item and the less likely it is that it will be guessed that saying No to it indicates neuroticism. The high negative correlation between N loadings and detection as an E Yes item shows that the higher the N loading the less likely it is to be guessed as an indicator of extraversion. The relationship between N loadings and detection as L items will be considered below.

The E factor loadings of the items are highly correlated with detection of the item as E Yes and highly negatively correlated with guesses that the item is E No. Thus as with neuroticism, the higher the factor loading on E the more likely the item is to be identified as an item indicative of extraversion. The correlations of E loadings with N guesses show that the higher the loading the less likely it is that the item will be judged to be a measure of N when answered Yes, and the more likely it is to be judged to be not a measure of N.

Thus, taking both scales together, it is clear that the factor loadings determine to a considerable extent (half or more of the variance) the detection of the item as being of a certain type.

Table 2. *Correlation between factor loadings and guesses that an item is of a certain type* ($n = 48$)

	N Yes	N No	E Yes	E No	L Yes	L No
N loading	0·76***	−0·53***	−0·66***	−0·07	−0·59***	[illegible]**
E loading	−0·33*	0·27*	0·71***	−0·75***	0·03	0·00

* $P < 0·05$. ** $P < 0·01$. *** $P < 0·001$.

Table 3. *Correlation between factor loadings and correct detection of an item as of a certain type*

(The correlations in bold type are between correct detections and the appropriate factor)

	N ($n = 24$)		E Yes ($n = 15$)		E No ($n = 9$)	
	F_1	F_2	F_1	F_2	F_1	F_2
N Yes	**0·03**	−0·50**	−0·44	−0·58*	0·27	0·35
N No	−0·40*	0·13	0·03	−0·26	−0·86**	0·56
E Yes	0·06	0·59**	−0·04	**0·73****	0·24	−0·20
E No	0·04	−0·28	−0·13	−0·09	0·04	**−0·76***
L Yes	0·38	−0·14	−0·46	−0·48	−0·42	0·80**
L No	−0·13	0·17	−0·40	−0·46	−0·25	0·77*

* $P < 0·05$. ** $P < 0·01$.

So far discussion of the items has been independent of the way they are arranged in the EPI as measures of E or N. The relationship between correct detection (correct being defined as a positive answer contributing of an EPI score) and factor loadings is now considered.

Correlations between factor loadings and correct detection as an N Yes, E Yes or E No item, and incorrect detection as N No, L Yes or L No are shown in Table 3.

The correlation between correct detection of N Yes items and the factor loadings of the items is effectively zero, but there is a moderate significant negative correlation between correctly detecting an item as N Yes and its factor loading on E. This means that although N factor loading differences between N items and detectability differences between these items are unrelated, the more an N item is a measure of E the less likely it is to be correctly judged as an N item. Also the negative correlation between N No and the neuroticism factor means that it is less likely to be judged as not measuring N if it has a high loading on N. Further, the correlation between loadings on the E factor and incorrect detection as an E Yes item shows that if the item is misclassified as E Yes it will in part be due to its loading on E.

The most important feature of the E Yes correlations is the high correlation between correct detection as E Yes and the factor loadings on E. This shows that even within the E scale items, over half of the variance in detecting an item as an E item is due to its loading on E. The other significant correlation in this case is a negative one between the E factor loadings and incorrect detection as an N Yes item. Thus if an E Yes item is misclassified as N Yes this will be due partially to its low loading on E.

As is the case with the E Yes items there is a considerable relationship between correctly detecting an item as E No and the factor loading on E. The relationship is negative since the factor loadings of the items were, of course, negative. The high negative relationship between factor loadings of these items on the N factor and incorrectly judging the items to be N No indicates that the more these items are measures of N, the less likely is it that they will be judged as N No items.

There is a considerable difference here with regard to the N and E scales: within the N scale the factor loadings do not determine correct detection, but they do within both the E scales. This contrasts with the correlations shown in Table 2, where there was a high correlation between the judging of an item as N and the factor loading on N. We believe this to be due to the greater factorial purity of the N scale. The mean N loading of the N items is 0·41 (S.D. = 0·08), whereas the mean E loading of the E Yes scale is 0·32 (S.D. = 0·16), and the mean E loading on E No is −0·24 (S.D. = 0·15). Since the mean of the loadings on N is higher than the E Yes and E No means, and the N scale standard deviation is half the standard deviations of the other two scales, then (1) each N item is a stronger measure of N, and (2) the dispersion is smaller, thus restricting the range. (It is also the case that some of the E items are better measures of N than E.)

The incorrect detection of N and E items as L scale items can now be considered. When the type of item is not considered (Table 2) a negative correlation is obtained between N loading and detection as an L Yes item and a positive correlation between judging it to be an L No item.

It would take up a considerable amount of space to go through the inventory item by item, but it is instructive to look at the three items of the N scale which were misjudged by almost half or more of the subjects as L No items. They are: 7, 'Do you sometimes sulk?'; 31, 'Are you touchy about some things?'; and 57, 'Do you often get "butterflies in your tummy" before an important occasion?'. Clearly many of the subjects thought that anyone who said No to these items was lying, rather than that saying Yes was an indication of neuroticism. The other item, an E Yes item, which was judged by 23 of the subjects to be a lie scale item, was item 27, 'Do you very much like good food?'. Many subjects considered an answer of No meant that the person was lying.

The other correlation of interest in connexion with misclassification of items as L scale items can be seen in Table 3. It shows that there is a considerable relationship between misidentifying E No items as L items and the factor loadings on E. Since these items had negative loadings the positive correlations show that the more negative the loading on E the less likely are the items to be misidentified as L scale items.

Conclusions

1. The data summarized in Table 1 show that if subjects are asked to detect items they can do this very adequately, virtually all items being detected with $P < 0{\cdot}001$. Thus if subjects so wish, they can correctly detect the items and present themselves as being more or less neurotic, extraverted and liars.

2. The personality of the subjects as assessed by the EPI, and their intelligence as assessed by the AH 5, do not determine their ability to detect EPI items, although it must be emphasized that university student subjects were used.

3. There is a high degree of relationship between factor loadings of all items and their detection as items of that type.

4. There is no relationship in the N items between differences in correct detection of items as belonging to the N scale and the difference in the factor loadings on N, but there is a high degree of relationship in the E items between correct detection of items as E scale items and factor loadings on E. This discrepancy is most likely due to the unsatisfactory nature of the E scale: the factor loadings on E are lower, have greater dispersion, and some are very low (less than 0·2).

5. We propose that this procedure, of having a group of subjects attempt to detect the nature of the items, be adopted as a preliminary precaution when a new personality scale is drawn up.

6. A paradox seems to exist, since the better an item (in terms of its factor loading) the more readily detectable it is. This suggests that inventories which are not susceptible to successful simulation will also be poor measures of the relevant variables. (Although the EPI was selected for this study other inventories such as the 16 PF and its derivatives are equally susceptible to simulation (Radcliffe, 1966).) If existing personality questionnaires are to be used when those completing them have something to gain, it will be necessary to examine further what people, under what conditions, can detect items. Or it is possible that a new approach will have to be adopted, that recommended by Meehl & Hathaway (1946), who have suggested in connexion with lie scales that: 'An effective use of a set of "subtle" items is only possible when the item pool is very large and the *initial selection* (not only the final validation) of items is ruthlessly empirical. Those items whose significance would not have been guessed by the test maker will then be equally mysterious to the testee' (p. 527). Clearly, this suggestion should be applied to *all* scales if there is to be any hope of their being effective.

The authors are grateful to Donald Sykes for his assistance in obtaining the data.

References

Eysenck, H. J. & Eysenck, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London: University of London Press.

Eysenck, H. J., Hendrickson, A. & Eysenck, S. B. G. (1969). The orthogonality of personality structure. In H. J. Eysenck & S. B. G. Eysenck (eds.), *Personality Structure and Measurement*. London: Routledge & Kegan Paul.

Heim, A. W. (1956). *Manual for the Group Test of High Grade Intelligence AH* 5. Slough: National Foundation for Educational Research.

Meehl, P. E. & Hathaway, S. R. (1946). The K factor as a suppressor variable in the MMPI. *J. appl. Psychol.* **30**, 525–564.

[illegible] R. P. (1968). Simulation of stable and neurotic [illegible] *Br. J. Psychol.* 59, [illegible]

[illegible] *Br. J. Psychol.* 60, [illegible]

[illegible] W. T. (1970). [illegible] test. *Br. J. Psychol.* 61, 91–93.

[illegible] EPI and MPI. *Aust. J. Psychol.* 18, 134–137.

(*Manuscript received 28 May 1970*)

Br. J. Psychol. (1971), 62, 3, pp. 403–407
Printed in Great Britain

CRITICAL NOTICE: REINFORCEMENT AND INFORMATION*

By D. E. BLACKMAN and G. V. THOMAS
Department of Psychology, University of Nottingham

The relationships between an organism's behaviour and its environment have been a traditional subject of interest in the development of psychological theory. However, the inadequacy has long been recognized of a simple mechanistic view that behaviour can be regarded as responses to antecedent stimuli. For example, it is seldom possible to predict an organism's behaviour given *only* a description of its contemporary environment. Skinner (1969), in discussing this problem, has suggested that 'no account of the interchange between organism and environment is complete until it includes the action of the environment after the response has been made' (p. 5).

One major system in which the consequences of behaviour are assigned a central role is that of reinforcement theory. The reinforcement principle is normally traced from Thorndike's statement of the law of effect. He suggested that a response may be selectively strengthened by the rewarding nature of its immediate consequences. These 'rewards' may be regarded as reinforcers, in current terms. Thorndike's original formulation emphasized both the effects of the reinforcer on motivational systems and also its temporal relation to behaviour. Two distinguishable positions have developed from these different aspects. The first has traditionally attempted to develop the motivational aspects of the principle. Perhaps the most familiar example of such an account is Hull's theory that a reinforcement reduces a central drive state of the organism. Skinner (1950), however, suggested that such theories of behaviour, which appeal 'to events taking place somewhere else, at some other level of observation, described in different terms and measured, if at all, in different dimensions', may deflect attention away from other important variables which are directly manipulable. The most important of such variables, in Skinner's interpretation of the reinforcement principle, are the relationships between events and behaviour (technically, the 'contingencies of reinforcement'). This may be regarded as a development of the second aspect of Thorndike's formulation, the temporal relationships between behaviour and its consequences. Further, a reinforcer is defined solely in terms of its *empirically* established effects on behaviour; no additional considerations are involved.

A recent book, *Feedback and Human Behaviour* by J. Annett, provides an attempt to compare the effectiveness of the concept of reinforcement with the concept of feedback 'information' in furthering our understanding of human skilled behaviour. The interpretation of the events which follow behaviour as feedback information has developed from cybernetics and machine control theory. The principal concept employed is that of a servo-mechanism, a system which regulates itself according to the consequences of its own output. In drawing an analogy between such machine systems and the biological system represented by a behaving organism, the consequences of behaviour are viewed as feedback which conveys information about a goal

* *Feedback and Human Behaviour.* By J. Annett. Harmondsworth: Penguin Books. 1969. Pp. 196. 30p.

or its attainment. As in cybernetics, this information is usually quantified in terms of Shannon & Weaver's information theory, or its developments.

In the preliminary chapter of his book, Annett discusses the general importance of the consequences of behaviour (described as 'knowledge of results'), in establishing and maintaining skilled performance. Subsequent chapters provide reviews of the effects of knowledge of results on motor skills (such as positioning, tracking, and reaction times) and perceptual skills (detection, monitoring, recognition, etc.). These reviews establish the suggestion that knowledge of results and the methods for delivering such knowledge play an important part in determining the efficiency of these types of behaviour. Chapter 4 is mainly about verbal learning (e.g. paired-associate learning, verbal conditioning, concept learning), but also includes maze learning and concludes with a brief discussion of programmed learning procedures. After these relatively empirical reviews, Annett turns to more general and theoretical matters. He examines interpretations of knowledge of results as 'incentive', 'reinforcement' and 'information', and concludes that the first two concepts refer to *motivational* aspects of knowledge of results, incentive as 'drive-inducing', and 'reinforcement' as 'drive-reducing' (p. 169). Annett has previously noted that 'the distinction between incentives and reinforcers is none too clear' (p. 36), and in his conclusions (ch. 8) his critique of motivational interpretations of knowledge of results is directed chiefly at reinforcement. He sees the central issue as the separation of the informational and motivational effects of knowledge of results (or as 'the reduction of one to the other'; p. 37). In his final chapter, Annett argues that the apparent motivational functions of knowledge of results can be accommodated within a strictly informational interpretation, and so favours the latter as the more parsimonious.

Annett himself summarizes his book in the generalization that 'the cybernetic analogy is more satisfactory and potentially more frutiful than the evolutionary analogy represented by reinforcement theory' (Preface, p. 12). Such a claim, if established, would clearly have important consequences for the development of psychology. One important feature of the cybernetic analogy put forward by Annett is the quantification it can apply to the stimulus situation in terms of 'information'. It is important to recognize at this point, however, that information ('in the information theory sense', Annett, p. 164) is not a psychological concept, but is strictly a *description* or quantification of a state of affairs in the environment (Attneave, 1959). This is not, of course, to limit the importance and usefulness of the concept – on the contrary. However, Annett attempts to discuss how information is *used* by organisms; in other words, to *relate* information to behaviour. It may be crucial to remember that 'the subject need not perceive objects in accordance with the experimenter's descriptive system' (Attneave, 1959, p. 83). So a distinction must be maintained between information as measured in the environment and the information *used* by the organism. Such a distinction is not made with any force by Annett, and as a result his treatment of the word 'information' may confuse.

Our particular concern in this review, however, is that Annett addresses himself to only one version of reinforcement theory, but dismisses the general concept of reinforcement in favour of information. Thus at several points, for example, he explicitly suggests that reinforcement ultimately depends on drive reduction (e.g. pp. 37, 133, 161). As we have noted earlier, this is not the only formulation of a reinforcement

principle Skinner's alternative is to define an event as a reinforcer when, and only when, it strengthens or maintains the behaviour to which it is related. It may be important in this context to note that the apparent circularity involved in defining an event as a reinforcer solely in terms of its 'empirically reinforcing properties' is not necessarily as serious a problem as Annett suggests (p. 162). For example, Hocutt (1967) has argued that 'one does not define a thing, one defines some property of a thing, or one defines a thing with respect to some classification' and 'one cannot define relational, as distinguished from intrinsic, properties of things independently of the relations these things bear to other things'. It could thus be argued that the property of being a reinforcer is a relational property which an event possesses *only* by virtue of standing in the relation of strengthening or maintaining *behaviour* (rather than reducing drives, for example).

This point assumes special importance when considering human behaviour, the subject of Annett's book, for it is not easy to sustain explanations of such behaviour in terms of ultimate drive reduction or even in terms of the reduction of secondary drives. It is only too clear from Annett's discussion that attempts to identify such drives are fraught with difficulties; but it may be important to recognize that many of these difficulties arise from the assumption of drive reduction and may not apply to Skinner's alternative definition of reinforcement.

The claim that distinctions should be drawn between the ways in which reinforcement has been interpreted has also been made by Keehn (1969) in a discussion of animal discrimination learning. It is our contention that Annett's discussion of reinforcement is effectively limited to what Keehn describes as S-O-R theories. But Keehn goes so far as to suggest that the alternative formation of reinforcement which we have earlier attributed to Skinner represents an alternative *paradigm* to S-O-R theories for the analysis of behaviour. This alternative explicitly excludes a dualistic conception of a behaving organism: Keehn has provocatively suggested that 'Although the symbol O is meant to emphasize *organismic* rather than *mentalistic* variables, the linear arrangement of the symbols serves to delineate a duality between organisms and their responses and to ossify behaviour in its traditional role as a mere exemplar of events occurring inside an organism' (p. 275). It may be disputed that Skinner's position (which is sometimes misrepresented) offers such a radical departure from traditional psychology as to justify its elevation to the status of an alternative paradigm as defined by Kuhn (1962). Nevertheless, it is true that this alternative interpretation of reinforcement is not adequately represented in Annett's book, although it is favoured by many current workers in preference to the motivational interpretation considered by Annett.

Feedback and Human Behaviour is published in a form which is likely to secure a wide readership; its ready availability and low purchase price are to be welcomed in a text which is designed for undergraduate use. In this context, it provides a valuable summary of much empirical data from the study of human skills. A further merit is that these data are discussed in the context of alternative theoretical positions. It is our opinion, however, that Annett may mislead the reader into a premature rejection of 'reinforcement' as an explanatory concept in psychology, because of his dissatisfaction with only one version of the reinforcement principle.

REFERENCES

ATTNEAVE, F. (1959). *Applications of Information Theory to Psychology*. New York: Holt, Rinehart & Winston.

HOCUTT, M. (1967). On the alleged circularity of Skinner's concept of stimulus. *Psychol. Rev.* 74, 530–532.

KEEHN, J. D. (1969). Consciousness, discrimination and the stimulus control of behaviour. In R. M. Gilbert & N. S. Sutherland (eds.), *Animal Discrimination Learning*. London: Academic Press.

KUHN, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

SKINNER, B. F. (1950). Are theories of learning necessary? *Psychol. Rev.* 57, 193–216.

SKINNER, B. F. (1969). *Contingencies of Reinforcement*. New York: Appleton-Century-Crofts.

(*Manuscript received* 1 *July* 1970)

REINFORCEMENT AND INFORMATION: A REPLY

BY JOHN ANNETT

Department of Psychology, University of Hull

The controversy over the various meanings of reinforcement has been going on for many years and it is not possible to retrace the tangled skein of argument in the short space permitted me. I would, however, like to clarify my position (implicit rather than explicit in the book) on the Skinnerian formulation of the reinforcement principle. It is true that Skinner does not get much space but his contribution to the field of human learning (programmed instruction included) is rather limited, consisting essentially of the repeated assertion that the effects of KR in human learning can be equated with the effects of reinforcement procedures in animal learning. My reformulation of the reinforcement *principle* (I oppose the continued use of the *term* since it has acquired so much excess meaning) encompasses the Skinnerian version. This version, defining 'an event as a reinforcer when, and only when, it strengthens or maintains the behaviour to which it is related' (Blackman & Thomas, above), is directly translatable into the statement: 'learning behaviour exemplifies a rule, namely, when the result of a response is of a certain type (say 'positive' without any motivational implication) then the tendency to repeat that response is increased'. It is convenient to say that the learner is 'using a rule' but no knowledge of unobservable processes is implied; this is purely an empirical generalization of the sort that Skinnerian philosophy recommends. This simple rule turns out to be empirically inadequate. For example, in the line-drawing experiment the learner's behaviour is better described by rules such as 'if the result of R 1 is "too long", then R 2 is made shorter and if the result is "too short" R 2 is made longer' (see pp. 165–7). I call such rules 'transformation rules' since they are the (empirically determined) rules by which the learner transforms feedback information from R 1 into instructions for making R 2. Human subjects, and probably lower orders too, exhibit a variety

of rules in their learning behaviour. In concept attainment, for example, a subject in a given set of responses may behave according to the rule 'if exemplar A was negative choose exemplar B such that it is different in only one attribute'.

My approach is entirely consistent with Skinnerian positivism in the sense that all such rules have the same empirical status as the 'reinforcement' rule. Indeed 'reinforcement' is but one of the class of transformation rules. I part company with conventional reinforcement theorists on the purely empirical grounds that the one rule does not fit all the observed facts. An obvious retort would be that to envisage an indefinitely large set of transformation rules infringes the law of parsimony. However, assigning one rule to a more general set is also parsimonious, especially in view of the failure of many serious attempts over the past 30–40 years to make do with only one. The more general notion of transformation rules in the cybernetic analogy seems, at this point in time, to offer a better way of bridging the gaps between simple and complex and between animal and human learning.

(*Manuscript received* 21 *September* 1970)

PAPERS TO APPEAR IN FORTHCOMING ISSUES

(Not previously listed)

[illegible], J. & [illegible], D. Office of Population, Censuses and Surveys, Social Survey Division, [illegible], London. A method for overcoming the problem of [illegible] in semantic differential research.

[illegible], J., [illegible], E. L. & HarAN, D. Department of Psychology, University of Manchester. A [illegible] of the [illegible] effect in [illegible] and [illegible].

[illegible], D. & [illegible], B. Department of Psychology, The George Washington University, Washington, D.C., U.S.A. Some findings relevant to the hypothesis that topological spatial [illegible] are [illegible] prior to Euclidean features during growth.

[illegible], F. E., [illegible], D. I. & Nickerson, C. J. Department of Psychology, University of [illegible], U.S.A. An evaluation of elicitation procedures for personal constructs.

[illegible], L. & Molander, B. Department of Psychology, Umeå University, Sweden. Verbal [illegible] effects in cross-modal transfer.

[illegible], L. Department of Psychology, University of Guelph, Ontario, Canada. On mental [illegible].

[illegible], J. S. Center for Advanced Study in Theoretical Psychology, University of Alberta, Edmonton, Canada. A dimension of mind rediscovered.

Kreitler, H. & Kreitler, S. Department of Psychology, Tel Aviv University, Israel. Cognitive orientation: a model of human behaviour.

Levy, P. Department of Psychology, University of Birmingham. Concept–scale interaction in semantic differential research: solutions in search of a problem.

Strongman, K. T., Wookey, P. E. & Remington, R. E. Department of Psychology, University of Exeter. Elation.

Taylor, J. G. Kitwe, Zambia. Underwater distortion: and plain distortion.

Wong, R. Department of Psychology, University of British Columbia, Vancouver, Canada. Infantile handling and associative processes in rats.

Br. J. Psychol. (1971), 62, 3, pp. 411–433
Printed in Great Britain

BOOK REVIEWS

Physiological Psychology. By PETER M. MILNER. New York and London: Holt, Rinehart & Winston. 1970. Pp. x+531. £6.00.

This book is probably the best general introductory textbook of physiological psychology available today. It is most suited to the needs of undergraduate specialists in psychology, but may also be of use to those with a more advanced knowledge of the subject. Physiological psychology is a rapidly expanding field, in which – as Milner points out – well over a thousand research reports are published every year. Some of these reports may radically alter or even revolutionize our theoretical understanding of certain problems. The difficulty for the student is to separate the wheat from the chaff. A good textbook makes this task easier by providing a clear basis of fundamental knowledge and by showing how recent experiments enlarge and alter the picture derived from earlier experimental work. Unfortunately, most textbooks fail in this second respect, either by offering a very limited and one-sided view of current research findings or by displaying an understanding of them which looks as if it was gained by reading abstracts and nothing more. Peter Milner has, on the whole, been successful in avoiding these stumbling blocks, not beleaguering the reader excessively with his theoretical biases and treating recent research in an illuminating and stimulating fashion.

The book is divided into six parts respectively entitled Foundations, Motor Mechanisms, Sensory Systems, Alerting Systems, Motivation and Emotion, and, finally, Memory and Learning. There is a 37-page reference section containing papers up to and including those published in 1969, and separate author and subject indexes. Each part is opened by a general introduction and every chapter is concluded by a summary, usually a page or more in length.

The first two chapters of Part 1 give a basic introduction to the structure of the neurone and of the human brain as a whole, whereas the third chapter introduces the most important techniques used by contemporary physiological psychologists. These chapters are models of limpidity; the second chapter in particular gives the concise kind of account of gross neuroanatomy that first-year students cry out for. Milner's explanation in terms of the phylogeny of the nervous system is copiously illustrated with diagrams and brain sections.

Motor systems are dealt with next because, in a sense, nervous systems exist to provide fine control of motor activity as it is on this that an animal's survival depends. Peripheral motor mechanisms are discussed before central ones because certain fundamental principles inferred from studying the former may help elucidate how the latter function. This part of the book serves as an introduction to the succeeding section on sensory systems and a later section on motivation – sensory processes being essential for motor control and there being no clear borderline between straightforward motor activity and motivated behaviour.

All the usual sensory modalities and aspects of sensory physiology are described in Part 3 and, in addition, there are passages summarizing research findings indicating the different role of the dominant and minor hemispheres in auditory and visual functioning. There are also chapters on the cortex and the neuropsychology of language. There are very few easily assimilable conclusions that can be drawn from work on the involvement of the brain with language and Milner's chapter on this topic is one of his least successful. It is liable to leave its reader with the impression of having covered too much too superficially.

The treatment of alerting mechanisms is adequate and one is made strongly aware of how jejune research on the physiology of attention has been until recently. However, the account of sleep mechanisms is confusing about Jouvet's biochemical studies, omits Oswald's contributions and skimps with those of Routtenberg.

One hundred and twenty-seven pages are allotted to the subjects of motivation and emotion. Regulatory and non-regulatory systems are described separately and summaries of our knowledge of self-stimulation studies and the function of the frontal neocortex follow. This is a considerable amount of material to compress into such a short space but the author does fair justice to his material with the possible exception of emotional behaviour, although one might cavil that for a book of this length the discussion of theoretical aspects of motivation is unnecessarily prolix.

The last section contains excellent brief accounts of the temporal lobe memory syndrome in

[illegible] and research on the electrophysiological concomitants of learning. Work on [illegible] hypotheses and 'molecularity' of memory is [illegible] presented [illegible] at present full of contradictory observations and impossible to describe without becoming too complex for most students' liking.

[illegible] textbook is concise, clear and well integrated and for this reason can be strongly recommended to students. The reviewer's [illegible] which the author freely admits may be seen by others as one of its greatest virtues.

ANDREW MAYES

Short-term Changes in Neural Activity and Behaviour. Edited by GABRIEL HORN and ROBERT A. HINDE. Cambridge University Press. 1970. Pp. vi + 628. £10 [illegible].

This volume contains the proceedings of an important conference held at Cambridge in [illegible] to review and describe recent research on changes in behaviour and the [illegible] which occur over intervals ranging from seconds to days. The main aim of the conference [illegible] relate the behavioural changes studied with neurophysiological phenomena at a variety of levels of analyses, e.g. neural and molecular. Much of the work reported requires the use of techniques which have only been refined in the last few years, such as the use of [illegible] to record from insect neurones over periods of several days described by Rowell.

The majority of the contributors are concerned with the physiological bases and the behavioural aspects of habituation. Habituation was used as the main example of a behavioural change for several reasons. First, it is a form of behaviour that can be precisely specified and reproduced under a wide range of conditions. Secondly, there are striking parametric similarities between habituation as it occurs in simple and complex systems and those habituation-like changes which take place at the neuronal level. Indeed, it seems probable that these neuronal changes provide the basis for the behavioural alterations. Thirdly, there have been considerable advances in our understanding of habituation in recent years. The 19 chapters of the book are divided into four sections with a brief introduction at the beginning of the first three. Section I deals with the behavioural problems associated with habituation. Section II goes on to discuss the physiological basis of this form of learning. Section III is concerned with the neural basis of learning and with learning processes other than habituation, and the final section is an epilogue, presented by Horn, in which the findings and conclusions of the conference are broadly reviewed.

In the opening section, Hinde first points out that habituation is a response decrement which occurs as a result of repeated stimulation within certain limitations and then delineates the common characteristics of this process. In contrast to Thompson and Spencer, he does not believe that sufficient evidence exists to suggest a unitary explanation of habituation in complex and simple systems. Kling and Stevenson are also sceptical about the possibility that habituation and extinction share the same underlying physiological mechanisms. It is difficult to make parametric comparisons between the two processes and such similarities as have been noted are too gross to provide the evidential basis for a unitary theory. In the final contribution of the section, Weiskrantz argues that information does not normally have to be in short-term memory before passing into the long-term system and that the former system may only be required for temporary holding of verbal information when access to the latter is blocked. Weiskrantz's argument is important, but if he is right, there is clearly no connexion between the short-term processes underlying habituation and those subserving short-term memory.

Most of the contributions to the second section consist in demonstrating that neural analogues of habituation exist. Thus, Segundo and Bell and Vinogradova show that neuronal changes within the vertebrate brain have the parametric characteristics of behavioural habituation, and Kandel, Bruner, Kehoe and others review work on the giant nerve cells of the marine gastropod *Aplysia* that is relevant to habituation.

The last main section reveals how little is known about the physiological and biochemical basis of behavioural changes that are more prolonged than habituation. It is argued by Bateson and by Horn, in the epilogue, that the chief problem is to establish that correlates of behavioural change are, in fact, the correlates of the specified change and not some concomitant. This problem is less severe for habituation than associative learning processes. However, Bureš and Burešová's attempt to classically condition the responses of single neurones in the brains of curarized rats may narrow this methodological gap. This section also contains a review of

[illegible] of imprinting.

[illegible]

A. MAYES

Brain and Early Behaviour. Edited by R. J. Robinson. London and New York, Academic Press. 1969. Pp. xvi + 374. £5.00.

Following the lead provided by ethology and [illegible] developmental psychologists [illegible] devoted a great deal of attention to the study of [illegible] and the effects of [illegible] forms of early experience on [illegible] development. [illegible] hundreds of papers and [illegible] the view of early experience [illegible] profound effects on subsequent development. Our view of the infant has [illegible] significantly over this period. No longer do we assume [illegible] uniform in their behaviour. However, despite [illegible] substantial increase in the amount of data available [illegible] great theoretical [illegible] achieved. To some extent this reflects conceptual difficulties in the develop[illegible]. It reflects also the lack of interaction between the various disciplines which [illegible] this important area. If a significant advance is to be made in our understanding of [illegible] factors which are involved in its control, then we badly need an integrated approach. This volume is one of the first serious attempts to bring different conceptual and empirical tools to bear on the problems in this field.

The book reports the proceedings of a study group which met to consider brain mechanisms and their relationship to early behavioural development. The group, made up of 15 scientists, was [illegible] in character, representing psychology, neurology, psychiatry and paediatrics. As the editor explains in his preface, the major portion of the group's time was given to discussions and this is reflected in the amount of discussion material and brief presentations included. The book is divided into a number of sections dealing with different aspects, each containing one or more contributions along with carefully edited but quite extensive discussion. In his introduction which sets the stage for what is to come, Prechtl raises a number of basic questions which run throughout the volume: what is development, what are the essential problems in studying early behaviour, what is known of the relationship between brain mechanisms and behaviour?

The three sections immediately following the introduction should be of particular interest to psychologists because they are concerned with the physiological basis of behaviour in the foetus and the newborn. Bergstrom develops a model of the functional growth of the brain in terms of a central core with high entropy and surrounding shells which are more highly organized. He suggests that in sequence ontogenetic development exhibits the different functional properties of the core and shell circuits. Oscillation between different ontogenetic states is seen in children and in certain pathologies. A particularly valuable paper by Humphrey considers the repetition of prenatal activity sequences in the neonate. Elegant and careful descriptions are given of avoiding movements of the head to stimulation in the perioral region, the finger movements which form the basis of grasping, and the development of plantar reflexes. Suggestions are made regarding the neuroanatomical basis of the postnatal repetition of foetal activity sequences. The paper is followed by an extensive and interesting discussion.

In a long paper Schulte and his colleagues from Gottingen deal with excitation, inhibition and impulse conduction in the spinal motoneurones of premature and full-term babies. The material on prematures is particularly interesting though the discussion on the effects of extra-uterine experience is surprisingly limited. Prechtl considers the importance of state in investigations of the behaviour of neonates and in so doing provides a good example of how physiological methods and behavioural observations may be profitably used together. His methods of data analysis and recordings are sophisticated and add much to the value of his work. In a study of an infant with holotelencephaly Wolff describes in detail the clinical course of this patient and links post-mortem findings to data on behavioural changes which occurred over the infant's four-month life span. Although the information provided by a single case is inevitably fragmentary, notions about the relation between functional competence and anatomical structure are stimulated by detailed

studies of this kind. Following a useful note on the nomenclature of stages of sleep by the editor two papers, by Petre-Quadens *et al.* and by Dittrichova, are devoted to the development of sleep patterns in the newborn.

The remaining 150 pages or so of the book are given over primarily to a consideration of aspects of behaviour in the neonate and through the first year. Two papers are devoted to studies on visual perception. Bronson traces the development of visual pattern recognition, which he believes appears in the second month, with Conel's anatomical evidence of marked neocortical development beginning at this period. The smiling response to a human face is used by Bronson as an index of pattern recognition, though he omits to mention [illegible] of smiling when a face is presented: do babies always smile when they recognize a pattern? Fantz's work which suggests that visual pattern discrimination occurs before two months (and hence presumably before neocortical maturation) is explained in terms of the organizational capacity of the bipolar cells at the retinal level. Bower describes briefly some of his experiments on perceptual constancies in babies which he investigates by using conditioning techniques. On the basis of these experiments he considers that true pattern perception appears much later than other perceptual functions and that the infant cannot handle multiple sources of information at the same time.

A very competent review of experimental studies of learning during the first year of life is given by Lipsitt. Most, if not all, of this research is based on the application of existing operant learning paradigms to neonates and young infants. One is left in no doubt that the infant can learn and that he is a much more competent learning organism than many believed him to be. It is a relief, however, when the author points to the need to relate questions about learning to the non-laboratory situations in which the baby normally lives and to which he is presumably well adapted; the spectre of volume upon volume of dreary investigations which characterizes much of the work on animal and human learning recedes. On the basis of his own studies of learning in babies Papoušek concludes that information processing ability is highly developed in the preverbal human infant and that an analysis of these basic forms of cognitive function may provide a valuable approach to specifically human ways of organizing behaviour. There is little attempt to relate data on infant learning to ideas of neural mechanisms which might form a substrate.

In an excellent paper on the development of early communication processes in the squirrel monkey (*Saimiri*) Ploog relates neuroanatomical data on brain maturation and neurophysiological data from brain stimulation studies to the emergence of complex mother–infant, adult–infant and peer communication processes. It is a little surprising that this is the only paper which deals with a non-human species but this is perhaps a reflexion of the paucity of investigations in the field.

In the latter part of the book two contributions are devoted to the problem of organization in behaviour: one is a general discussion, the other a contributed paper. This is probably the weakest section of the book and with the wisdom of hindsight it is clear that an invited paper should have been devoted to this complex and difficult problem. Finally, three brief but valuable contributions consider structure–function relationships.

An excellent summary is provided by Wolff who considers the group's successes and failures in the light of what they set out to achieve. It is an insightful evaluation which draws out the common themes that recur throughout the discussions. One implicit assumption is very firmly shaken: interdisciplinary meetings do not in some mystical way give rise to a chain reaction of cross fertilization. As Wolff states clearly it is futile for the student of behaviour to assume that someone else will solve all his neuroanatomical or computer problems. Clearly there is no substitute for a reunification of the sciences and where better to begin this than in psychology. A special word of congratulation must go to the editor. Much of this book is based on free-roaming discussion and the overall cohesion of the volume is a tribute to his unobtrusive skill in welding together a complex mass of material. The book is a valuable contribution in an extremely important area and deserves to be very widely read.

KEVIN CONNOLLY

Contemporary Problems in Perception. Edited by A. T. Welford and L. Houssiadas. London: Taylor & Francis. 1970. Pp. 175. £3.00.

[illegible] of papers originated at a N.A.T.O. Advanced Study [illegible] held at the University of [illegible] in July 1968. First published in [illegible] (Vol. 13, No. 1, 1970) [illegible] volume. The editors [illegible] some of the current [illegible] of research on human perception which [illegible] to be important in the future on both theoretical and practical grounds'. The contributions [illegible] mainly from Northern Europe and from the English-speaking [illegible] a considerable variety of interests and methods to give a [illegible] picture of the contemporary scene.

The papers fall into three groups. The first group describes theoretical approaches. A. T. Welford discusses the principles of perceptual [illegible]. He leads the reader [illegible] from the traditional terminology of [illegible] to the current [illegible] and templates. As a piece of introductory writing this could [illegible] R. L. Gregory continues with a more detailed study of [illegible] perceptual [illegible] to computer principles, leaving us with a tantalizing glimpse of what might have been if only Russ Hanson had not been killed, tragically, in his private plane. Finally, two papers offer a close study of special problems. [illegible] describes an '[illegible]' [illegible] of [illegible] discrimination, and Kabrisky and his colleagues use anatomical findings and computer analogues to formulate hypotheses about pattern recognition.

The second group comprises experimental studies. F. Metelli contributes a particularly neat study of transparency. L. Houssiadas reports a Michotte-type examination of perceptual components, though he leaves us in doubt whether the same subjects were used throughout, so that special 'sets' might have been established. R. B. Freeman examines visual slant and shape. M. Farne demonstrates visual after-effects in the third dimension which cannot be reduced to those on the frontal plane. M. Fatouros and J. Glick find that whiteness constancy is rather better in adults where the comparison objects are of similar rather than contrasted shape. Finally, D. M. Cuceloglu compares three cultures – American, Japanese and Turkish – in judging the facial expression of emotion. He finds main dimensions of pleasantness, irritation and non-receptivity, and concludes that 'there seems to be a facial code employed in the communication of affective meaning which is to a great extent, although not wholly, common to different cultures'.

The third group examines various aspects of the relation between perception and action. A. F. Sanders is concerned with the changes in perceptual strategy which accompany tasks involving different shifts of attention. R. S. Easterby makes an illuminating application of gestalt-type principles to the design of symbols for machine displays. W. M. Smith observes that not only does perception guide behaviour, but behaviour guides perception; and proceeds to a vigorous and incisive attack on the problems which this raises. His thoughtful discussion of temporal integration, of delayed visual information, and of retinal inversion made this the outstanding paper of the book for the present reviewer. He concludes: 'the chief task before us is not to concentrate upon new and different variations of the kinds of experiments many of us have done to study temporally and spatially displaced feedback, but rather to turn our major effort towards trying to conceptualize this whole problem area in a new way – in a manner which conforms more to a cybernetic model'.

Smith's conclusion might well be generalized to perceptual studies as a whole: the re-examination of concepts is currently more important than the continued accumulation of experimental findings of the traditional kind. This need is illustrated by a striking inconsistency between the theoretical and experimental sections of this book. The theoretical papers are all variations on a theme which is stated by Welford as follows: 'Perception is not to be understood as a matter of sensory data alone, but rather in terms of more or less elaborate constructs or frameworks.' Yet the editors describe the experimental papers, quite correctly, as 'aimed at accounting, in terms of stimuli reaching the eye, for features of perception which are often regarded as matters of inference and judgement...'. This divorce between theoretical intention and experimental practice might be described as *the* contemporary problem in perception. It seems to have slipped unnoticed through N.A.T.O.'s net.

R. B. JOYNSON

Theories of Cognitive Consistency: a Sourcebook. Edited by ROBERT P. ABELSON, ELLIOT ARONSON, WILLIAM J. MCGUIRE, THEODORE M. NEWCOMB, MILTON J. ROSENBERG and PERCY H. TANNENBAUM. Chicago: Rand McNally, 1968. Pp. xvii + 901. $17.50.

According to Leon Festinger, rats and men come to love the things for which they have suffered or exerted themselves. Merely to lift this massive volume by hand requires a considerable physical exertion, and anyone who has carried it around for weeks and wrestled with its 84 chapters of close argumentation, densely illustrated empirical controversy and often turgid prose should, according to cognitive dissonance theory, have become deeply infatuated with its charms. My own feeling, however, is that it is impossible to form any generalized attitude towards such a wide-ranging diversity of material, presented by no fewer than 84 different contributors.

Apart from a handful of masochistic reviewers, not many people are ever likely to read the book from cover to cover. Undergraduates will no doubt occasionally be referred to Section I, 'Consistency Theories: Six Positions', in which each of the editors undertook to present a systematic review of his own approach in the field of cognitive consistency. These six chapters turn out to contain rather candid and speculative informal discussions of the more controversial aspects of the theory to which each of the editors adheres, usually illustrated by modestly selective reviews of the empirical evidence on a few outstanding problems. Balance, congruity and dissonance theories are dealt with at a high level by Newcomb, Tannenbaum and Aronson respectively, Rosenberg elaborates on aspects of his affective-cognitive consistency theory, Abelson takes steps toward an integration of symbolic psycho-logic with dissonance theory, and McGuire ruminates on his now largely abandoned work on the structure of human thought. None of these chapters constitutes anything like a systematic review of a theoretical position, and none is really rudimentary enough for undergraduate reading. They are all stimulating and enormously valuable to readers with some background knowledge of the field, but they provide no substitute for such background knowledge and they lack comprehensiveness. As the editors were well aware, competent systematic reviews are available elsewhere, and the reader is referred to the 'standard' reviews of the literature on balance, congruity and dissonance at the beginning of each of the relevant chapters.

Section II is devoted to ten excellent and unique chapters on the relationship of cognitive consistency theories to other psychological theories. These include a chapter by Sarnoff on psychoanalytic theory and cognitive dissonance, two very different behaviourist reactions to consistency theories by Collins and Bem respectively, a revealing exegesis of Heider's original balance model in the larger framework of gestalt theory by Jordan, and a résumé and response from a consistency viewpoint by McGuire. This section contains a critical examination of the fundamental assumption of all consistency theories, that people are motivated to seek cognitive consistency or at least to avoid inconsistency. A later subsection of the book, entitled 'Assumptions about Cognitive Structure', is devoted specifically to this question, but some of the most interesting arguments are to be found in Section II, and these two sections are not cross-referenced. This is but one example of the inevitable leakiness of many of the compartments which make up the formal structure of the book.

The final two sections, Section III and Section IV, deal with 'preconditions and properties' of inconsistency, and 'responses and consequences' respectively. Each is divided into about half a dozen subsections, covering such topics as 'Self and the Consistency Process', 'Inconsistency and Psychological Stress' and 'Commitment as a Mediating Variable'. The titles of these sections and subsections should not, however, be read too literally. All five papers in a subsection grandly entitled 'The Temporal Course of Inconsistency Reduction' are in fact confined to summaries of research on post-decisional dissonance reduction only. A subsection entitled 'Selective Exposure to Information', containing four excellent and sharply contrasting evaluations of the current status of the selective exposure postulate of cognitive dissonance theory, is misleadingly classified under 'responses and consequences'. Selective exposure is in fact a means of avoiding dissonance arousal rather than a response or a consequence.

The style and quality of the 84 chapters are, of course, uneven. Some chapters are refreshingly original, some provide extremely useful discussions of areas of research, and some seem either redundant or irrelevant. There are, for example, two chapters, by Watts and Silverman respec-

[illegible] 'almost unnecessary', but such redundancy is even truer of some of the other contributions.

[illegible] may well form the basis of emergent areas of research for years to come. The sections on 'Individual Differences in Reactions to Inconsistency' and 'Modes of Resolution' in particular will probably play an important role in providing a springboard for future research [illegible]. Other parts will become out of date rather rapidly because of the extraordinary pace of research in some areas of cognitive consistency. The discussion in Section II and elsewhere of Bem's interpersonal replications of dissonance experiments has already been made obsolete by the work of Jones and his colleagues, published since the book went to press.

The editors were probably unduly optimistic in hoping to cover so much ground with grass growing so rapidly underfoot. The expansiveness of the field is due very largely to the phenomenal growth of research on cognitive dissonance, which began stealing the thunder from the other consistency theories a mere decade before this book was published. Despite numerous attempted refutations, dissonance theory appears to be more talked about than ever: Festinger receives the greatest number of citations in the name index, and 'cognitive dissonance' the greatest number in the subject index. This understandable imbalance is unfortunately purchased at the cost of neglect of some other theories. Although the list of references, boldly headed 'Bibliography', contains well over a thousand entries, there are several, and possibly many, important gaps in the coverage of the book. Nowhere, for example, is Cartwright and Harary's signed digraph theory given systematic treatment, and Feather's balance model—frequently discussed in contemporary literature on attitude change—is not even mentioned in passing.

The purpose of the book as originally conceived by its editors was to provide a theoretical synthesis of different approaches in the field of cognitive consistency, but 'the issues involved in "cognitive consistency" turned out to be much deeper and broader than we had imagined', so the attempt at synthesis was abandoned. As a result, although it contains a great deal of value, this book lacks a unifying purpose.

ANDREW M. COLMAN

Mathematical Psychology: an Elementary Introduction. By CLYDE H. COOMBS, ROBYN M. DAWES and AMOS TVERSKY. Englewood Cliffs, N.J.: Prentice-Hall. 1970. Pp. xi+419. $10.95.

Mathematical Psychology: an Elementary Introduction is the best introductory text available for those who wish to understand many of the quantitative solutions that have been suggested for psychological problems. The special virtues of this book are its wide coverage and its readable style. There are sections devoted to scaling, decision, learning theory and information theory, as well as a useful mathematical appendix.

It should be noted that Coombs, Dawes and Tversky have directed their book at those with an appreciation for psychological problems. Thus, in keeping with the chapters on scaling, the locus of 'elementary' in the book's title would be relatively high on a scale of 'usage of *elementary* by textbook writers'. However, the reader need not be especially sophisticated in mathematics, nor interested in all of mathematical psychology, since the chapters are relatively self-contained.

The authors' style is to abstract the important concepts from each particular mathematical development that they consider. The non-mathematically inclined reader often finds his intuitive understanding increased by the fascinating examples that illustrate almost every abstraction. The mathematically sophisticated reader cannot fail to enjoy these examples, although he may complain about the lack of problems, or the lack of detail about some of the more difficult proofs and methods. All of these factors contribute to *Mathematical Psychology*'s usefulness as an introduction for psychologists.

The first section of the book is about scaling. 'Measurement', the *theoretical* aspect of the topic according to Coombs *et al.*, and 'scaling', the *empirical* aspect, are discussed. Included as

examples of measurement are relatively simple descriptions of semiorders and additive conjoint measurement. Psychologists wishing to understand these types of scales, in order to apply them to a specific problem, will find these introductory descriptions invaluable.

In the chapter devoted to empirical aspects of scaling, Thurstone's contributions and scalogram techniques are topics. In addition, metric distance and multidimensional scaling are mentioned briefly. As an example of the book's style, the detailed description of Thurstone's law of comparative judgement is followed by a presentation of some data obtained by Coombs in which Thurstone's 1927 scale for the heinousness of crimes is compared with a similar 1966 scale. Because the comparisons are interesting in their own right, they make immediately clear the usefulness of scaling.

A chapter on non-numerical models complements the more conventional topics in scaling. Graph theory, set theory (applied to group relationships), and computer simulation are discussed. The General Problem Solver programme is described in relative detail as an example of simulation.

The second section of the book focuses on some topics in choice behaviour. The growth and use of utility theory is emphasized. Brief mention is also given to some specialized utility theories, such as the motivation-incentive model, as well as to the work on subjective probabilities contributed by the Bayesian theorists. Luce's choice axioms and a random utility model of Coombs's (based on unfolding theory) are then described, under the heading of choice where subjects are uncertain about their preferences.

Game theory is discussed separately. Both strictly and partly competitive games are described and some data on bargaining agreements are given.

Included with these chapters on choice behaviour and game theory is signal detection theory. A concise description is given of this theory with some emphasis on questions of sensory thresholds, and the success with which the theory has predicted human performance.

The last section of the book is about learning and information theory. The part on learning has two chapters: one introductory and the other presenting theoretical and empirical developments of the last 20 years. The introductory chapter begins with the concept of a learning curve and ends with a description of Markov chains. The mathematical concepts are illustrated by the prisoner's dilemma problem. Emphasis is on both theoretical predictions and empirical questions of interest, e.g. hyperaggressive boys are compared to normals.

The chronological development of mathematical learning theory is then considered. The major emphasis is on a classification of the many operator and state models that have been put forward in the last two decades. Probability learning and some simple concept identification situations are discussed to provide an empirical basis for these models.

The last chapter of the book, on information theory, emphasizes differences between statistical and psychological usage of terms such as information, transmission, and redundancy. A classification and description of experiments is then given.

In summary, *Mathematical Psychology* is a survey of a wide variety of topics, in which information is presented in a style understandable to most psychologists. The classifications the authors have imposed on many topics are useful, and often as interesting as the examples that have been chosen as illustrations. Thus the book is recommended to those wishing to gain an introductory understanding of the many psychology problems to which mathematics has been applied.

C. WEIR

Minnesota Symposia on Child Psychology. Volume 4. Edited by JOHN P. HILL. London: Oxford University Press. 1970. Pp. ix + 275. £2.40.

Volume 4 continues the tradition of diversity and immediacy of this series. On the basis of work with animal subjects, Ader conceptualizes the effects of early experience on susceptibility to disease; Dolhinow, also working with animals, breaks at last into that bastion of ignorance 'play'; Crandall and Devereux, in widely differing programmes, have important comments to make about the child-rearing antecedents of achievement and peer-group orientation respectively. Herbert Pick, taking a wide perspective, attempts to make some sense out of perceptual development, and Rohwer studies the effects of imaginative conceptual activity upon learning.

The concept of stress in psychosomatic research is Ader's particular concern: why, when exposed to the same environmental pathogens, some individuals manifest disease and others do not. The general concept of 'stress' he rejects as not particularly useful, as various stressors

do not behave in the same way (not an altogether surprising conclusion, one would think). The apparently reasonable hypothesis, however, that early experience might permanently affect the functioning of the adrenocortical system is examined and found inadequate. Much potentially interesting material here is lost by a limited vocabulary and a repetitious presentation. Careful editing would have added to its readability and impact.

In contrast to Ader's laboratory approach is Dolhinow and Bishop's paper on the play of free-ranging non-human primates. Here is detailed, expert, acute observation in the field at its best and it is reassuring to find that straight observation is not too *démodé* for inclusion in a book of this kind. It provides a useful, thought-provoking and timely contribution to the literature on that absorbing activity of the infant, play. The material is set well in context, and familiar and some unfamiliar ideas about play are examined in the light of the reported data. So little is known about this topic that no one paper could hope to provide all the answers; this one, however, sets the interested research worker firmly in the right direction.

In a longitudinal study of achievement development using Fels's data, Crandall and Battle tried to find the earlier behavioural precursors of motivational orientations and achievement behaviours in three areas: intellectual, vocational and academic achievement. These distinctions were made on a 'hunch' basis initially – a 'hunch' which proved fruitful, as evidenced, for example, in the discussion of the relationship between maternal independence training and achievement effort, where only male intellectual achievement seems to relate to the maternal variable. This is a useful tidying-up operation in a field which was becoming untidy.

Devereux's concern with the role of peer group experience in moral development is with preadolescents. He argues the need for an optimum level of support and control in the home. His 'optimum levels' theory predicts a curvilinear relationship between almost any parent practice variable and almost any child behaviour outcome. If this theory holds, it could account for the consistently low and confusing correlations found in researches on parent–child behaviour, as most of them have been based on simple linear models. The finding that both high permissive homes and high punitive homes produced peer-orientated children he cites as part of the evidence for his theory.

In considering the role of mental elaboration and proficient learning Rohwer uses the technique of learning word pairs. He finds the nature of the elaborative context does make a difference in the efficiency of learning. Further work is required here for a fuller understanding of elaboration phenomena.

I think it unfortunate that Pick's paper on spatial orientation does not detail his actual instructions to his subjects, since from past experiments it is clear that the instructions can distort the data. Pick is broadly concerned with whether there are qualitative changes in perceptual development as striking as the changes Piaget has identified in the development of cognition. He finds little evidence for this. In his search for some coherence he uses the concept of a system – a concept which, in the terms in which he uses it, is useful for the whole of developmental study. Simple processes add together to make complex processes and the result may be that uninteresting quantitative changes, taken in combination with others in an integrated system, might then add up.

This series certainly 'adds up' to a significant contribution to the literature of ongoing research in child psychology.

MARY CROXEN

Fundamental Statistics for Psychology. By ROBERT B. MCCALL. New York and London: Harcourt, Brace & World. 1970. Pp. viii + 419. £4.30.

Introduction to Statistics: Selected Procedures for the Behavioural Sciences. By ROBERT FRIED. New York and London: Oxford University Press. 1969. Pp. viii + 304. £2.60.

General Statistics. By AUDREY HABER and RICHARD P. RUNYON. Reading, Mass. and London: Addison-Wesley. 1969. Pp. xiii + 321. £3.50.

The difficulties experienced by non-mathematical students in basic statistics courses are formidable and notorious. It is not surprising, therefore, that each of these three books offers a different diagnosis and treatment.

Fried asserts that courses fail because they are 'exercises in obscure symbol manipulation of hypothetical data from mythical populations'. He claims to write as a user of statistics addressing other users. It is strange, then, to find that his text contains little real data and little practical advice; he does not even distinguish between statistical and practical significance.

Haber and Runyon believe that courses fail because they are 'dull and uninteresting gauntlets that test the stamina, tenacity and frustration tolerance of the student'. They aim at a readable and instructive text by using real data, newspaper extracts and relevant humourous extracts from Daryl Duff's *How to lie with Statistics*, by eliminating coded scores and correlation charts, by 'trying to preserve the distinction between population parameters and sample statistics', by 'reversing the trend to relegate descriptive statistics to a position of secondary importance', by including operating instructions in each statistical table and by making end-of-chapter exercises an integral part of the course.

McCall believes courses fail because they ignore 'the principles of the psychology of learning'. He does not indicate which psychological learning theory he has in mind and his remedy turns out to be no more than a few unoriginal pedagogic devices which owe more to common sense than to psychology, viz. important definitions are set off in distinctive type, mathematical proofs are given in panels isolated from the main text, no steps are omitted in algebraic proofs or derivations and each step is explained, statistical tables carry operating instructions, symbols are accompanied several times by their verbal names and concepts and formulae are briefly reviewed as needed.

Although all these innovations may help, they are clearly too trivial to have more than a marginal effect and certainly have little chance of achieving the effects claimed by their authors. To find sugared water offered as elixir of life is not unusual, but to find it in the academic market-place is disturbing.

In content the three texts cover roughly the same ground, with McCall's text, as befits its greater size and price, giving the widest, deepest and most detailed coverage. All include chapters on frequency distributions, measures of central tendency, variability and relative standing, regression and correlation, a short introduction to probability, simple tests of significance, chi-square tests of contingency data, a brief introduction to analysis of variance and a quick look at non-parametric methods. The main differences lie in the order of the chapters and in arbitrary inclusions and omissions of specific statistics. Fried also adds a bonus chapter on time-domain distributions illustrated by data on galvanic skin responses and cardiac responses.

What is surprising is that none of the authors presents an up-to-date version of basic statistics. None treats conditional probability, Bayes's theorem or decision theory; only McCall shows any influence of the latter and then only in relation to classic hypothesis testing. Non-parametric methods are treated as appendages to the main theme and relegated to short sections at the end. In Fried and Haber–Runyon the distinction between parametric and non-parametric statistics is confusedly given and appears to be imperfectly understood by these authors. Thus neither includes measurement requirements in their list of parametric assumptions and they appear to recommend parametric statistics for use with data, such as ratings, which fail to meet the requirements of metric scales. Consequently, important non-parametric statistics such as Kendall's tau are ignored. Indeed, apart from a quick bow in the direction of the distinction between nominal, ordinal and metric data, none of the texts shows any influence of measurement and scaling theory.

In the treatment of analysis of variance all three ignore repeated measures designs, types of models (except for a brief mention by McCall), comparisons of means (except, again, McCall who gives only the long-discarded, invalid method of separate t tests), tests of additivity and tests of linearity. This failure is perhaps most disturbing in the Haber–Runyon text, as these authors expressly claim to 'reflect the latest statistical advances and technological changes' and assert that they 'have not hesitated to introduce new statistical techniques'. The latter is limited to the use of Sandler's A statistic as a computationally simpler version of the t test for correlated samples.

Even without these shortcomings it would be difficult to recommend the Fried text as it contains too many mistakes and, in several places, notably when dealing with tests of significance, is confused and confusing. The Haber–Runyon text also fails to live up to its claims. Despite several useful innovations, it offers little that is new and omits much that is normally provided in a basic text. Nor does it succeed in its objective of restoring descriptive statistics to a position of primary importance; the coverage of this field differs little from the other texts.

The McCall text is easily the best of the three for psychologists. It is the least out of date, has the fewest omissions and is lucid and well produced. It provides a useful introduction to this field.

A. B. ROYSE

Language and Psychology: Historical Aspects of Psycholinguistics. By ARTHUR L. BLUMENTHAL. New York: Wiley. 1970. Pp. x + 248. £3.75; paper, £2.25.

This book traces the development of psycholinguistic theories from Wundt down to the current burgeoning of linguistic theorizing and concomitant psychological research. The author's thesis is that American psychologists are unaware of a period of intense concentration by linguists and psychologists in the last quarter of the 19th century on problems concerning language which have recently become fashionable again. This early tradition was swept aside by behaviourism and it was not until 1950 that psychologists and linguists rediscovered each other; simultaneously new movements arose in linguistics apparently not directly related to their intellectual predecessors, which were soon galvanizing psychological research. Like the other volumes in the 'Perspectives in Psychology' series, the book consists of selected readings and linking commentary in about equal amounts. The readings offered, several of them new translations by the author, certainly show a striking modernity in 19th-century thinking, with notions such as the primacy of the sentence, abstract relational structures and the creative nature of children's language learning being advanced. It seems to the reviewer, however, that no evidence is offered that these notions depended on 'a very close collaboration between early experimental psychologists and linguists'. In fact, the period of this collaboration is unrepresented in the selection, which opens with the psychologist Wundt asserting his right against the linguist Paul to theorize concerning language. Those ideas of Wundt which are closest to modern linguistics were in fact opposed to those of influential contemporary linguists and not derived from them.

The author claims to have restricted himself to the history and development of psycholinguistic theory in the present volume and to have reserved discussion of early experimental work for a paper which was unfortunately not yet available at the time of this review. Nevertheless, the present volume does include chapters covering investigations of 'Language Acquisition' and 'The Psychology of Reading', so the dichotomy is not strictly observed. This is fortunate, as it is difficult to understand theories adequately without some knowledge of the experimentation of the period, and it is questionable whether the 'considerable experimental and observational data' produced by Wundt and his students should have been omitted. The reviewer was left wondering about the nature of these data, and was unable to discover more from any readily available source. How far was the decline in Wundtian psycholinguistics due to its failure to generate means of investigating the issues Wundt discussed theoretically, and Wundt's own belief that such a programme would not prove fruitful, in addition to the dualism which rendered it abhorrent to behaviourists? Perhaps Wundt's own creation of experimental psychology helped extinguish psycholinguistics.

The selection of readings provided for the period covered is reasonably representative, so far as can be judged by one who shares the unfamiliarity with much of the literature of those for whom the book was devised. Modern theorists are represented (inevitably!) by Lashley, Chomsky, Miller and Lenneberg. The author is to be congratulated on making readily available material which has been neglected. The book provides an interesting case history in the birth, life and death of ideas and should provide some counter to the widespread assumption that 'recent' is necessarily synonymous with 'more important' in psychology.

J. M. WILDING

Studies in Thought and Language. Edited by JOSEPH L. COWAN. Tucson, Arizona: University of Arizona Press. 1970. Pp. 228. $6.50.

Studies in Thought and Language, edited by J. L. Cowan, is a collection of six papers, five of which were presented at a symposium organized by the Department of Philosophy at the University of Arizona in February 1968. Cowan introduces the collection as an example of 'interaction among the fields of philosophy, psychology and linguistics', and as a useful corrective to attitudes of mutual distrust which have prevailed between these disciplines in the

recent past. Of the papers included, two are by psychologists (Mandler and Osgood), and the remainder are by linguists or philosophers (Vendler, Wells, Cowan, and Ziff).

Mandler's paper, entitled 'Words, Lists, and Categories', does not concern itself with the philosophical and linguistic questions which are treated by the other contributors, but reviews some recent experimental work on the topic of the organization of memory for words. Organization is defined in terms of relations among the words in a person's vocabulary, and Mandler argues that the main types of relation involve seriation, categorization or relational [illegible]. A number of experiments on seriation and categorization are described. The aim of the paper is perhaps to argue that laboratory studies of list learning have implications for a general theory of word memory, and this view is supported by some studies of 'free emission' of words (e.g. giving animal names).

A long paper by Osgood, called 'Interpersonal Verbs and Interpersonal Behaviour', gives a detailed exposition of an extended programme of psycholinguistic research which Osgood started in 1964 in collaboration with Kenneth Forster, and includes some cross-cultural studies. The work was based on 'an approach to the measurement of meaning which employs the rules of usage of words in combination as a means of discovering the semantic features of the words thus combined'. Osgood's account treats meaning as 'simultaneous bundles of semantic features', although the possibility of translating this notion into the language of mediational S-R theory is distinctly underplayed in this paper. There is also only a brief discussion of the semantic differential as a device for measurement of meaning. The studies reported by Osgood started from intuitive listings of features of interpersonal verbs, and developed into empirical tests of the acceptability of word combinations among samples of language users. The paper describes various possible approaches to analysis of this type of data, and concludes with a tentative taxonomy of possible organization for a semantic system.

It is interesting to compare Osgood's approach to meaning with Vendler's paper, entitled 'Say what you think'. Vendler follows Austin and other 'ordinary language' philosophers in proposing that thought and language can be studied in the context of a classification scheme for illocutionary acts. He argues that Austin's analysis can be extended to take account of 'the various kinds of verb-objects illocutionary verbs take' in order to develop grammatical distinctions between verbs of thinking and saying. For a psychologist who is interested in language and cognition a major problem seems to be one of determining the relationship between descriptions of competence of the type given here by Vendler (and also by Osgood) and theories which can properly be called psychological. The remaining papers in the collection give some assistance on this point. Wells writes about the distinction between *a priori* and empirical questions in relation to theories about the relationship between thought and language. He sees one concern of philosophy as being to distinguish 'the meaningful from the meaningless, and within the meaningful, the *a priori* possible from the *a priori* impossible'. An empirical science, such as psychology, may then work 'within the *a priori* possible to distinguish the factually true from the factually false'. Cowan's paper concerns the same question, with special reference to the impact of the work of Chomsky and Katz on psychology. Cowan argues that transformational grammar has been handed to psychologists wrapped in a mentalistic myth which constitutes a source of avoidable confusion: in particular, such grammars, however excellent, 'are not themselves psychological theories and still less philosophies of mind', but 'constructions of the imagination wrought by grammarians'. Finally, a short paper by Ziff on the topic of 'Understanding' attempts to define the distinction between two people who hear an utterance, but where one understands while the other does not. It is encouraging for cognitive psychology that Ziff concludes that this distinction cannot be adequately represented by descriptions of appropriate reactions to the utterance, or of dispositions to respond appropriately: rather, the notion of understanding appears to imply the occurrence of 'analytical data processing of some sort', although the nature of this processing remains mysterious.

In summary, *Studies in Thought and Language* contains an interesting and provocative set of papers, which might be profitable reading for psychologists interested in cognition and language.

P. H. K. SEYMOUR

The Psychology of Speech and Language: an Introduction to Psycholinguistics. By Joseph DeVito. New York: Random House. 1970. Pp. xii + [illegible]. $[illegible].

One of this book's virtues is the clarity with which it sets out the author's views and his method of organizing the material. The book is intended as an introduction to psycholinguistics for beginning students of speech, linguistics and psychology. The first section consists of an introductory chapter on the nature and function of language, followed by an account of three theoretical approaches: linguistic theory, learning theory and communication theory. The second half of the book describes the application of psycholinguistics to four particular research areas: speech and language acquisition, breakdown, differences and effects.

The theoretical chapters cover a great variety of theories, including several that may be [illegible] to an experimentally orientated psychologist. Examples are Alfred Korzybski's theory of general semantics, which offers a set of [illegible] devices for ensuring [illegible] with reality, Walker Gibson's classification of styles as tough, sweet or stuffy, and McLuhan's 'the message is the medium'.

While it is undoubtedly desirable that students' horizons should be widened, the drawback is that there is less coherence in the treatment of those [illegible] and those theories which have played a central role in the development of psycholinguistics. The accounts given of [illegible] and S-R models, information theory and generative grammars are brief and oversimplified to an extent which is sometimes misleading. A striking example is the difference between language and speech. Having carefully defined the former as 'the abstract system which is manifested or actualized as speech, as vocal utterances', in much of the discussion that follows DeVito proceeds to ignore the distinction entirely. For instance, in the middle of an account of linguistic methods of analysing language he introduces a description of the type-token ratio, with no indication that this is a measure of particular samples of speech or written utterances. It is perhaps a little odd, too, to find in this same section on 'Word Analysis' a discussion of kinesics, dealing with communication through bodily motions, which is obviously one of the author's special interests.

A further point is that in spite of commending theories for their research potential, DeVito gives no general account of the types of experiments that have been stimulated by learning theory, information theory and linguistics. Of the four chapters on applied research areas only the one on acquisition attempts a systematic comparison of the different theoretical approaches. The chapter on 'Speech and Language Breakdown' concentrates on aphasia and stuttering, and even within these topics on a limited selection of theoretical concepts and research methods. 'Speech and Language Differences' draws together differences in individual styles, dialect variation, the effect of social context (including an investigation into the difference between suicide notes and ordinary letters), linguistic relativity and, to end with, the biggest difference of all, that between human and animal languages. The curiously titled final chapter on 'Speech and Language Effects', subtitled 'Rhetorical Psycholinguistics', turns out to deal exclusively with balance theories of attitude change such as Osgood's principle of congruity and Festinger's cognitive dissonance, no special attention being focused on linguistic variables.

To sum up, a psychology student reading this book would certainly be introduced to a wide range of terms used in the field, which he might be able to follow up in the quite useful suggestions for further reading. But he would be unlikely to gain any clear idea of the relative importance of the various theories, the way in which they relate to basic issues in the study of language or their implications for current psycholinguistic research.

JUDITH GREEN

On Intelligence. The Toronto Symposium on Intelligence: 1969. Edited by W. B. Dockrell. London: Methuen. 1970. Pp. 267. Price £3.00.

It is all too easy for symposia to consist of papers which, although relating to a broad general theme, remain very esoteric and cover narrow areas of inquiry. Such a criticism cannot be levelled at this symposium, which clearly satisfies its aim of examining both basic theory and educational practice in the light of recent research.

The perennial, basic issues of innate and acquired influences, cultural differences, general and specific abilities and hierarchical models are linked to the practical problems of test con-

struction and educational practice. Burt, Vernon and Jensen give lucid accounts of their theoretical viewpoint. Burt also replies to some of the criticism of his concept of inborn general mental ability and Vernon includes a succinct account of his cross-cultural investigations. Jensen's paper is almost a monograph in itself, providing a widely based rationale for his hierarchical model of human abilities and a summary of the experimental evidence from which the theory was derived, together with more recent studies to test specific hypotheses.

Although Guilford was not present at the conference, his model of intellectual functioning received considerable attention, both in a paper by Merrifield on 'Structuring Mental Acts' and in the discussions which followed each session. Piaget was also well represented. Tuddenham described his attempt to convert Piaget's concepts of intellectual development into test items conforming to accepted psychometric criteria. This paper led to a useful discussion of such questions as the discontinuity of development, cross-cultural studies and other political and theoretical problems relating to Piagetian concepts. The work of Ausubel and Gagné is also discussed.

Theoretical issues and experimental findings were related to problems of educational practice in a paper by Evans on 'Intelligence Transfer and Problem Solving'. In this he claims to have isolated a problem-solving factor from the analysis of a longitudinal study of mathematical performance. This paper provoked discussion on the broad question of teaching problem-solving strategies. Warburton's paper on the British Intelligence Scale provides a detailed and instructive account of problems met in constructing a scale of intellectual abilities which is related to current theoretical accounts of intellectual functioning. It is appropriate that the symposium should be dedicated posthumously to Professor Warburton, for his paper reflected the success of the conference in presenting broad conceptual and theoretical issues in the context of practical problems such as those of test construction and remedial programmes. This symposium is essential reading for anyone concerned with the assessment and development of intellectual abilities and the clear and concise style of most of the papers makes it a useful source book for undergraduates seeking information on particular theories and techniques. B. AKHURST

Festschrift for B. F. Skinner. Edited by P. B. DEWS. New York: Appleton–Century–Crofts. 1970. Pp. x + 413. £3.65.

As a *Festschrift*, this book is thoroughly disappointing. In dedicating it to B. F. Skinner on his 65th birthday, the intention was, one assumes, to support the claim made by the editor in the first sentence of his preface that Skinner is 'one of the great men of our times'. One looks, then, for the reappearance of some of Skinner's more influential or controversial papers; for appraisals of his contribution to psychology from his colleagues, both advocates and critics; for authoritative reviews of the present state of those areas of research which have been most affected by his thinking. But one looks in vain in this book for such a general contribution to the psychological literature, and one is left with the strongest feeling that an opportunity has been sadly missed.

The first part of the book is not entirely misdirected. It opens with the paper by Skinner, which was first published in Boring and Lindzey's *A History of Psychology in Autobiograph*, vol. v (Appleton–Century–Crofts, 1967, pp. 385–413). Surprisingly, the source of this paper is not adequately referenced; nor does it appear in the otherwise useful bibliography of Skinner's publications which follows it. Thereafter, two interesting, but inconsequential, reminiscences are to be found, the first of the period 1926–31 at Harvard by Keller, and the second, by Ferster, of the pigeon laboratory there in the 1950s. Neither of these papers takes one by storm; although they are pleasant enough, they cast little new light on Skinner the man, or on Skinner the psychologist.

It is from page 47 onwards that this *Festschrift* loses its direction. Faced with the admittedly difficult problems of who should be invited to contribute papers of scientific merit, and of how such contributions should be edited, decisions were effectively avoided. In short, no invitations were issued. Instead, 'many people were informed of the plans for a Festschrift', and were told, it seems, that manuscripts published in the *Journal of the Experimental Analysis of Behavior* in 1969 which incorporated a dedicatory footnote might be included in the book. The result is 'a highly selected more or less random collection of papers' (editor's preface, p. x), or, to paraphrase, an arbitrary assortment of reports written within the conventions of a research journal

in order to communicate with a like-minded readership rather than as contributions to a *Festschrift*. Each paper is printed exactly as it appeared in the *Journal*, and each contains its dedicatory footnote (variously worded – a mildly interesting study in itself). Each concludes with the date on which it was received (by the *Journal*, of course). The only obvious changes in format are the replacement of running titles by the title of the book, and, perversely, the removal of any clues as to the exact location of each paper within the *Journal*.

Of course, such a selection does give some indication of the present state of one research area which owes a very great deal to Skinner. However, a more accurate indication of this area may be achieved merely by perusing the *Journal* which contains all but one of these papers and many others too. Some of the papers which appeared in volume 12 but which are not included in the *Festschrift* seem of more significance and general interest than many which do appear in the book. But they did not contain the dedicatory footnote; perhaps one is allowed to wonder why this should be. To be fair, some of the papers reprinted in the book deserve their place by any standards. In particular, the more theoretical papers at the back of the collection are both important and interesting. For example, Herrnstein's ('On the Law of Effect', pp. 377–400) has already achieved a detectable influence on current research. This appeared in the *Journal* in 1970. MacCorquodale's retrospective appreciation of Skinner's *Verbal Behavior* (pp. 340–350) provides a timely and sensitive defence of some of the more abused aspects of that work. But here too one feels a certain unease, for this same author published in the very *Journal* a critique of Chomsky's famous review of *Verbal Behavior* which might seem even more appropriate for a volume dedicated to Skinner; this paper, in fact, appeared in volume 13 of the *Journal*, but nevertheless before Herrnstein's paper noted above. The final paper in the *Festschrift* is by Bijou ('What Psychology Has to Offer Education – Now', pp. 401–407). This was first published in the *Journal of Applied Behavior Analysis*, but was before that an invited address to the Division of School Psychologists at the 1968 A.P.A. Convention. Bijou has consistently and patiently argued for the application of Skinnerian principles to education, but this is not the most forceful example of his work, perhaps because of its original forum. However, this paper does emphasize the sad fact that there is no other paper in the *Festschrift* on applications of Skinner's principles to human behaviour outside the laboratory. Not even the burgeoning 'behaviour modification' techniques obtain a mention.

It will be seen that further comments about this *Festschrift* might easily become a critique of the editorial policy of the *Journal of the Experimental Analysis of Behavior*, which would hardly seem appropriate here. In the absence of any linking commentary in the *Festschrift*, one can either evaluate each paper separately in this way, or hope to isolate areas of potential growth and importance. The latter would seem inadvisable if based merely on this arbitrary selection of papers. But there is perhaps one problem of general importance which deserves brief attention. Skinner has consistently argued for controlled research procedures using individual subjects rather than statistical groups, an argument which has been incorporated into a policy decision by the *Journal*. When this research strategy is used with care, it is convincing and powerful; but when it is used badly, it becomes very bad indeed. When recently reviewing Hendry's book *Conditioned Reinforcement* in the *Journal of the Experimental Analysis of Behavior*, Gollub (*ibid*, 1970, **14**, 361–372) expressed his reservations about the adequacy of some of the experiments reported in that book. Such reservations could be equally appropriate in evaluating some of the individual reports in this *Festschrift.* To take one example, there has recently been increasing research interest in a phenomenon where animals (usually monkeys) consistently emit operant responses which are periodically followed by intense shock. This so-called response-produced shock phenomenon seems paradoxical in motivational terms, but is the sort of phenomenon which experimental analysts of behaviour have always thought imporant; it directs attention to the experimental histories of the animals rather than to the assumed motivational forces prompting their 'paradoxical' behaviour. One does not need too fertile an imagination to envisage some abnormal patterns of human behaviour as being related to such experimental findings. The phenomenon could therefore be important both theoretically and practically. It is, then, appropriate that three papers in the *Festschrift* (pp. 201–240) should report experiments investigating 'response-produced shock'. Yet these papers all lead one to the belief that the basic phenomenon has yet to be adequately defined in a well-designed experiment. It is by no means clear what are the appropriate controlling variables; control conditions are noticeably lacking, and in one of these experiments in particular, there is a cavalier disregard for systematic changes

in independent variables. This is a most unsatisfactory state of affairs, and it seems sad to find it emphasized in a volume intended to honour Skinner.

This book adds almost nothing to the evaluation of Skinner's role in the development of psychology. There is little indication of the heated controversy which his writings have prompted in the widest sphere, and no impression of how such controversies have affected other psychologists for good or ill. The book fails in its purpose of honouring the man and his work, and one hopes that a more adequate volume of considered appraisal will mark a future happy occasion.

DEREK BLACKMAN

The Dynamics of Action. By JOHN W. ATKINSON and DAVID BIRCH. New York and London: Wiley. 1970. Pp. xi + 380. £5.25.

Kurt Lewin used to insist on the principle of adopting Galilean modes of thought in psychology, in place of the Aristotelian modes he saw as prevalant. J. W. Atkinson enthusiastically accepted Lewin's principle and has tried to implement it conceptually, theoretically, and empirically for several years. In the present book, Atkinson and Birch would seem to want to make the leap from a Galilean to a full fledged Newtonian presentation: *The Dynamics of Action* is an attempt to construct a systematic and integrated theoretical psychology of motivation, resting on a very few basic assumptions manipulated with mathematical precision. The book, although highly mathematical, is no empty formalism. 'We did not take an abstract mathematical system and use it as a representation of motivation... Instead, we have taken our ideas about motivational processes, phrased them in the language of mathematics, and used mathematical operations to derive consequences' (p. viii). This is, of course, what Newton did too.

The basic theoretical assumption, reasonably enough, is that of inertia: 'A behavioral tendency, once it has been aroused, will persist in its present state until acted on by some force that either increases or decreases its strength' (p. 10). The implication is explicitly drawn that the organism must be considered as constantly active. Early versions of the assumption are credited to Freud and Lewin. Therefore, say the authors, the basic problem of motivation is to describe the 'change in the activity of an individual, from some initial activity that is already in progress when the interval of observation begins to another activity later.... The two classic problems of motivation, persistence of an activity and initiation of an activity, are two sides of the same behavioral coin – a change in activity' (p. 23). Throughout the book, the determinants of the change from one activity to another are described in terms of an application of the principle of change of activity: 'the time taken from some arbitrary initial point when activity *A* is occurring until the final point when activity *B* replaces activity *A* is directly proportionate to the difference between the strength of the tendency sustaining activity *A* at this final point in time and the initial strength of the tendency to undertake activity *B* and is inversely proportionate to the magnitude of the instigating force for activity *B*' (p. 22; symbols deleted).

The first chapter explains the conceptual scheme and introduces the mathematical treatment of the concepts. Chapter 2 introduces the idea of behavioural families, and discusses substitution, displacement, and the common fate of family tendencies. Cognitive activity is briefly considered, and is held to be described by the same dynamics as other behaviour is. The authors favour a view of cognitive or perceptual (but not necessarily conscious) activity as a necessary mediator of motor activity. Chapters 3 and 4 consider the thorny question of how one is justified in making inferences from observed behaviour to behaviour tendencies and back to behaviour again, both in the simplified ideal case and in the real world of the laboratory and the natural environment. The authors point out that behaviour is rarely unitary, and show that any prediction of behaviour based on single instigating forces (such as specific drive level) will be systematically misled. These chapters contain several implicit criticisms of S-R approaches. Chapter 5 interprets classical and instrumental conditioning in terms of the authors' schema. Chapter 6 deals with thought – 'the cognitive correlates of an instigating force'. Cognitive expectancies are considered as summating components of the 'total instigating force' to engage in a particular activity. Individual differences in personality are related to differential sensitivities to various components of a family of instigating forces. The treatment so far has been solely in terms of the *instigation* of behaviour. Chapters 7 and 8 deal with the complications involved in a consideration of punishment, failure and inhibition, and lead the authors to a theory of *resistance* to action. Both behavioural and cognitive effects are considered, both from a contem-

[illegible] and a [illegible] learning standpoint. The [illegible] ninth chapter attempts to com- [illegible] to determine the resultant [illegible] [illegible] the authors [illegible] the [illegible] that has been undertaken to test their theory, and [illegible]. [illegible] are defined in a glossary.

This is a [illegible] book, but a good one. The authors emphasize the tentative nature of many of their [illegible] and often present alternative [illegible]. The [illegible] mathematical [illegible] of the theory will deter some readers, but the [illegible] of the theory [illegible] as possible. It is an open question whether this degree of mathematical [illegible] is at present at any rate desirable or justifiable in a general psychological theory. The authors [illegible] that it is, and explicitly set out to test their contention by presenting the theory. In [illegible], generality and applicability, the theory sets an impressive standard of [illegible]. The [illegible] question left, as the authors acknowledge, is whether it has much substantive [illegible] and that will be for the future to decide.

B. D. MACKENZIE

A History of Scientific Psychology: Its Origins and Philosophical Background. By D. B. KLEIN. London: Routledge & Kegan Paul. 1970. Pp. 907. £8.00.

This [illegible] and expensive book is a British edition of a work published earlier last year by Basic Books in the United States. The author, Emeritus Professor D. B. Klein, was formerly a professor of psychology in the University of Southern California, and may possibly be known to some older British readers as the author of a much better than average book on *Mental Hygiene*, published in 1944.

The present work is a stimulating and scholarly history, extremely well done if it is taken on its own terms. It is emphatically *not* a history of scientific psychology, and the title can only be described as misleading. It stops at Wundt, just when psychology really got launched as a science, and the whole of the 19th century is sketchily treated apart from certain key figures, who are picked out for special exemplification and treatment—namely Herbart, Lotze, Bain and Wundt. The whole background of scientific development in the biological sciences in the 19th century gets only passing reference, and men who certainly played a significant part in the development of a scientific psychology are barely mentioned. Galton gets half a line, and Spencer is once incidentally named. This is in no sense a complete or rounded history of psychology's origins.

The book starts with the Greeks, and is essentially an account of the philosophical precursors of psychology. It is a psychologist's history, not a historian's, nor a philosopher's. Professor Klein is interested primarily in showing the relevance of the philosophical past to the psychological present. He is constantly jumping ahead to draw parallels and point morals. Thus in the chapter on the early medieval period the account of Plotinus merges into a discussion of William James's views on mysticism, and the account of Occam is linked to the canon of Lloyd Morgan. This is interesting and suggestive to psychologists, even if it is not history as such. Klein's comparisons jump the ages and transcend their contextual backgrounds; he is constantly looking for anticipations, and the bearing of past events on contemporary issues, all of which is a bit upsetting to the historian, who would like to see a firmer sense of the general cultural and ideological setting of the ideas and the people discussed.

Nevertheless, within its chosen frame of reference Klein's book has many merits. It is based on first-hand reading of much of the source material, and a good many of the recent secondary sources. Klein is well versed in contemporary views on the nature of science, the nature of history and the philosophy of mind. After reviewing the Grecian and medieval background of psychology, he interposes a whole long section of over one hundred pages on history, the mind-body problem, and problems of selfhood and volitional freedom, before he plunges into a discussion of Hobbes, Descartes, Locke, Spinoza, Leibniz, Kant, Berkeley, Hume, the Scottish school, the associationists and the Mills. There is a liveliness and breadth about Klein's writing which is fitting in an author who often refers to, and has obviously been much influenced by, William James. Above all, Klein's book is an interesting one, interesting to the ordinary student of psychology, who is not concerned with the minutiae of historical scholarshp, but who is enriched by seeing contemporary issues against a broad background of long-standing philo-

sophical problems. The book shows all the marks of having been a labour of love, and the author conveys something of this feeling to his readers. Most students of psychology would benefit greatly from reading it.

L. S. HEARNSHAW

Behavioural Worlds: the Study of Single Cases. By P. G. HERBST. London: Tavistock Publications. 1970. Pp. xi + 248. £2.50.

The author of this book firmly holds certain opinions. Depending upon your persuasion, you will either find them unacceptable (because they threaten your theoretical and methodological investments), or find that they are opinions which may be entertained as talking points, or, should you already be one of the converted, find them convincing. These opinions are: that the postulates of physical science are not generally applicable to behavioural science; that one cannot assume, for the behavioural world, that there are measuring scales which will remain invariant regardless of the object being measured and invariant over time; and that every behaviour system (person or group) has its unique laws and operates according to its own measuring scales. Under these circumstances, behavioural science would seem to be for optimists but Herbst is not deterred and proceeds to explicate a basis for the study of single cases. If you accept his postulates, this is the only remaining possibility.

The task of behavioural science, for Herbst as it was for Lewin, is to find appropriate conceptual representations and to develop methods adequate for the discovery of behavioural laws. It is unlikely, Herbst considers, that these laws will be as invariant as those which hold, for example, for gases. If they were, we should have been able to discover them by 'population-sampling' of behavioural processes using invariant measuring scales. Instead, we find that behavioural universes evolve their own laws and measurement scales (pp. xi, 20) and further, that 'if there is a conclusion to be drawn from half a century of population-sampling study techniques it is that the existence of [laws like the gas law] is highly improbable as far as the behavioural sciences are concerned' (p. 6).

With the exception of a homily on deception in laboratory experiments (pp. 12–13), some mysticism (pp. 53–55) concluding with the assertion that 'the non-subjective world... lies beyond the boundaries of science', and some entertaining speculation on the relationships between the culinary arts, good living, and medical practice (pp. 209–210) the book is concerned, in the main (parts I and II) with Herbst's longitudinal and cross-sectional case study techniques applied to his coal-mine data and to pupil–task relationships.

Herbst argues the case for the difference between behavioural and physical laws and extends the argument to the proposition that behavioural laws can best be discovered through the analysis of single cases. Case studies, we know, should be concerned with the analysis of the interrelationships of the many variables affecting the case rather than, as in population-sampling techniques, with a few variables affecting many cases. Unfortunately for Herbst, his chosen phase-space approach restricts consideration to a few variables – affecting single cases. This would, perhaps, be defensible if the chosen variables could be shown to be especially important ones but on this matter, particularly with regard to the pupil–task study, he is more casual than convincing.

More contentious are his assertions about 'parametric steady states'. First, he tells us that such conditions must exist before we set about measuring the variables of the particular behavioural system but he does not tell us about the criteria for steady states. He sidesteps this issue by talking about 'initial learning phases' (p. 86). Secondly, there is a strong hint that the steadiness of the parameters is to be assessed by the steadiness of the variables themselves, involving circularity. Allow variables to vary too much and 'critical limits, beyond which changes in the value of parameters are induced' (p. 60), might be exceeded.

Herbst needs parametric steady states just as much as animal psychologists need to know the history and condition of their animals. One must be able to assume some degree of contextual invariance before lawfulness can be revealed but one must also specify these conditions and the independent means by which they can be assessed.

Herbst is quite correct in saying that his cross-sectional case study technique is an alternative to his longitudinal technique. In a sense, they are identical. In the longitudinal case, given parametric steady state conditions, the sequence of events (i.e. time) can be disregarded (p. 59). One simply has to deal with a sample of states of the system. In the cross-sectional case, providing

the different *sections* (school subjects, in his pupil-task study) are sufficient, we again have a sample of states of the system – but more economically. The crux of the matter, however, is the adequacy of the sampling. Make the conditions of the sampling similar enough (the steady states steady enough or the sections close enough) and lawfulness is bound to obtain. This is Herbst's Moby Dick. But it is also the problem of representativeness, experimental control, pure case, validity, and so on.

Part III (chapters 10–14) consists of essays only very loosely related to the rest of the book. Chapters 10 and 11 take another look at the group life-space problem which interested Lewin, chapter 12 applies behaviour system theory to Heider's point of view on cognitive balance, the speculations on sense modalities and developmental psychology in chapter 13 should have been omitted, and chapter 14 (already published) concerns a linear transition model and organizational commitment. The value of the book is in its revitalizing of persistent methodological problems which refuse to vanish when we look the other way.

Incidentally, unwary developmental psychologists should not be trapped into thinking that Herbst uses 'longitudinal' and 'cross-sectional' conventionally. As we have seen, time is to be disregarded. His longitudinal method is not really longitudinal at all.

R. E. C. PENNY

LSD, Marihuana, Yoga and Hypnosis. By **THEODORE X. BARBER.** Chicago: Aldine. 1970. Pp. xiv + 337. $8.95.

Dr Barber is well known for his extensive writings on the subjects of hypnotism and yoga, and it is he, probably more than anyone else, who has been responsible for taking a good deal of the mystique and sensationalism out of what he himself refers to as the 'so-called soft areas of psychology', presenting them as susceptible to the same techniques of analysis as are applied to other, less dramatic phenomena. I commenced with considerable enthusiasm to read this, his latest book, expecting to find within its pages both an extension of the author's earlier ideas and some new light thrown upon these very interesting issues. It was therefore with some disappointment that I discovered that only chapters 1 and 2, which deal with psychoactive drugs, were completely new, the remaining five being revisions of papers already published in various journals.

In another direction too the book failed to live up to expectation: the title seemed to hold the promise of an attempted integration of several apparently disconnected topics, but if anything the reverse in fact proved to be the case. Not only is the book sectionalized into two parts, the psychoactive drugs being dealt with quite separately from yoga and hypnotism, but even within each section the fragmentation of the subject-matter is further emphasized by the appearance of lists of references at the end of each chapter rather than being collected together at the end of the book. The unity of the book is clearly intended to lie in the author's expressed aim to 'illustrate a way of thinking and a method of analysis' of the topics with which he deals, and it is at this level that the work has to be judged.

The methodological approach which Dr Barber advocates consists in deciding, from the volume of published literature and by means of new investigations, which antecedent conditions are causally related to the observed effects of any given treatment. The specification of stable antecedent–consequent relationships will, he suggests, provide the foundations of empirically based theoretical formulations. It is, however, only in the matter of hypnosis that the author comes close to reaching this goal, and even in this one wonders whether the rather ponderous end-result justifies the great labour involved in its attainment.

Nevertheless, the book is, in places, most enjoyable. There are fascinating accounts of physiological changes allegedly produced under hypnosis, from the curing of warts to the control of pain, and Dr Barber deals fully and critically with such claims. The short chapter on hypnotic age regression is a model of critical argument.

The weakest part of the book is undoubtedly the section dealing with drugs. There is a quite unaccountable lack of discussion of work on the behavioural effects of the drugs on animals, without which no empirically based theory of drug mechanisms can be considered complete. On the other hand, there are detailed expositions of drug effects upon chromosomes in tissue culture and upon embryogenesis in rats and mice, most of which seem to be largely irrelevant to the stated aim of theory construction within a psychological framework. There is also a distinct impression on occasion that the thread of argument becomes lost in extensive digressions into

[illegible] dependence upon marihuana.

There are further surprising omissions and superficialities in other parts of the book. I would have thought that the great volume of work of recent years on [illegible] would have merited more than a passing mention in any discussion of [illegible].

Despite these shortcomings and an abundance of typographical errors and [illegible] spacings between letters in some of the tables, I have no doubt that the book will [illegible] to the bookshelves of all who are interested in hypnosis. Dr Barber's reputation [illegible] least that. I suspect, however, that its major function will be that of a reference [illegible] it will undoubtedly fill admirably but which is, none the less, somewhat removed from the author's original intention.

F. N. JOHNSON

Institutionalism and Schizophrenia: a Comparative Study of Three Mental Hospitals 1960–1968. By J. K. WING and G. W. BROWN. Cambridge University Press. 1970. Pp. 260. £3.75.

Following the mental hospital reforms of the 1950s there was an urgent need for evaluation of the effects of the reforms, especially on the patient who seemed destined to spend long years in hospital. The reforms had clearly loosened up entry to the hospitals but could they be shown to be affecting those who somehow did not manage to be discharged? Indeed were such patients (largely schizophrenic) subject to some sort of intrapsychic process which was quite independent of the environment and so not amenable to a more therapeutic environment? The authors set out to answer these questions by cross-sectional analyses of a carefully drawn sample of chronic women patients and their physical and social environment on three occasions in an eight-year period. This book is the result of their study and a considerable achievement of persistence, scholarship, and cooperation.

The rating and interview techniques were all well validated and reliable, and the emphasis of the authors is upon quantifying the areas they observed. There will be some objection that this inevitably results in a somewhat arid and abstract picture of the people and situations they describe but they have included enough material to communicate the ethics of the three hospitals—at least to a professional audience. One interesting writing technique is to put all the detailed statistical data at the end of the text, which certainly helps the continuity of the prose.

In their very good review of disease and the social environment, the authors state that it is their intention to find causes and not just associations. This is a critical decision on their part and it is possible that their method does not really allow them ultimately to argue causally since they were not only dealing with a multivariate situation but also were not attempting to observe any of the processes actually at work. In the event the consistent evidence of the relationship between clinical conditions and social environment is so great that they *can* justifiably argue a causal effect—but their method does not allow them to demonstrate it as such. Despite the considerable complexity of the situation studied (three hospitals at two or three different points in time) their results are remarkably consistent so that they conclude that 'a substantial proportion, though by no means all, of the morbidity shown by long-stay schizophrenic patients in mental hospital is a product of their environment' after having allowed for all the alternative theories such as dissimilar selection of the patient in the three hospitals and the so-called 'burning out' process. They add one most important finding which is that the actual process of reform (which ameliorated conditions in both hospital and patients) may also have a dynamic of its own and ultimately cease. In the most advanced hospital there was clear evidence of a deterioration in both patients and hospital—albeit still with a high range of comparative achievement.

Various aspects of their findings will also variously please. Drug treatment appears to be of little account and they underscore the well-known but ill-recognized fact that some hospitals were demonstrably more therapeutic a considerable time before the ubiquitous chlorpromazine was available. They also conclude that the most important single factor associated with handicap

[illegible] and wound was [illegible] with which it was associated.

The main contributions of this study seem to be twofold. First, that it demonstrates with great [illegible] the importance of [illegible] framework not only in treatment but also in the [illegible]. This is the main [illegible]. Secondly, that [illegible]. The present [illegible] away from the old [illegible] treatment [illegible] to general hospitals and community care [illegible].

One criticism—the subject index is terrible!

D. F. HOOPER

Modern Trends in Psychological Medicine—2. Edited by JOHN HARDING PRICE. London: Butterworths. 1970. Pp. xi + 381. £5.60.

This second volume, although not wholly dissimilar to the first in terms of the topics that have been included, generally reflects the developments that have occurred in understanding, treatment and research techniques over the last 20 years. These are exemplified in the chapter by Penrose in which he provides a brief but useful review of cytological findings relating to impairments of mental ability and to behavioural disorders. Autosomal and sex chromosomal aberrations are discussed with an emphasis upon the clinical aspects of the associated syndromes.

Szafran reports a sophisticated study of the effects of ageing on professional pilots utilizing physiological and signal detection techniques. Although a little out of place in this volume, it is an interesting contribution and Szafran offers the tentative general conclusion that the adequacy of brain function in a complex skill is dependent on the efficiency of the lung-heart system rather than upon age *per se*.

Oswald presents an excellent review of EEG sleep studies with special reference to insomnia, dreaming and drugs. He makes the point that the rebound effect in paradoxical sleep after withdrawal of amphetamine or medium acting barbiturates may be detected for two months. The implications for clinical research should not pass without note.

Schizophrenia is the subject of two chapters, one concerned with aetiology and the other with treatment and outcome. Tanimukai and Himwich review the evidence for a biochemical aberration in the schizophrenias with an emphasis upon the indoleamine hypothesis and conclude that if such an aberration exists it is quantitative rather than qualitative. Mandelbrote's examination of treatment and outcome in schizophrenia is notable mainly for omitting consideration of the large amount of recent work using operant conditioning techniques for modifying the behaviour of schizophrenic patients.

Cooper and Shepherd illustrate the ecological approach in examining the deleterious effects upon psychological functioning of war, socio-cultural change and biological change. A comprehensive discussion of anorexia nervosa is provided by Russell, who considers symptomatology and aetiology in detail. The qualitative differentiation between reactive and endogenous depression is the focus of the section by Kerr and Roth. Their consideration of psychological and biochemical processes in aetiology, in addition to the problems of classification, will be useful to clinicians and research workers.

Price describes recent developments in the organization of psychiatric services, including rehabilitation and adolescent units, and emphasizes the role of the general community. The chapter by Stafford-Clark on supportive psychotherapy, when compared with that by Whitehorn on the same topic in the 1948 edition, suggests that there has been little development in this field over the past 20 years.

Brady summarizes the principal techniques of behaviour therapy but omits description of recent innovations such as implosion and modelling. The temporal lobe is the subject of three separate contributions by Corsellis, Driver and Falconer and Taylor. Respectively, they examine pathological anatomy, electroencephalographic diagnosis and epilepsy. Each is an excellent review in depth of techniques and findings and taken together they represent a valuable source of information and evaluation.

Not unusually, in a book of this kind, there are topics which could have been omitted and others

which should have been included, such as psychosomatics and recent developments in interoceptive conditioning. However, this will be a useful volume, both directly and indirectly, for those whose fields of interest have been covered, for there are 1309 listed references.

A. T. CARR

Man Machine Simulation Models. Psychosocial and Performance Interaction. By ARTHUR I. SIEGEL and J. JAY WOLF. New York and London: Wiley-Interscience. 1969. Pp. xiv + 177. £4.75.

Computer simulation is becoming a fashionable development in the mechanization of the social sciences. In recent years, for example, there have been quite a few published reports of simulations in the general area of organizational behaviour. The present book is a how-to-do-it account of several exploratory digital computer simulations of task behaviour at individual and group levels. In particular, it describes the development of two simulation models: a unitary (i.e. one man) or dual operator model and a group model. The unitary dual model is exemplified by tasks such as landing an aircraft on an aircraft carrier, launching an air-to-air missile, in-flight refuelling of aircraft and air interception; the group model is applied to a number of representative tasks of a nuclear submarine crew on long-term cruising alert. Some idea of the complexity of the group model simulation may be obtained by listing some of the included variables: crew members' task specialisms and levels of proficiency, members' personality orientations, morale and cohesiveness, crew size, communication structures, objectives data, nature and state of equipment. The idea essentially is to build up a viable model of a socio-technical system in order to predict man–machine effectiveness.

The book can be read on two levels. First, as a detailed description of the complex logic of digital simulation. As such, the minutely spelt out mathematical and logical presentations make difficult, and at times mortifying, reading but this of course is intrinsic to the subject matter. The laboured comprehension does at least lead to an appreciation of how the complex data are actually organized for computer processing. Secondly, the book can be read in terms of the verisimilitude of its simulations and the extent to which the simulations reflect accepted psychological theory. It seems to me that several basic assumptions are made on the psychological variables that are overly limiting. For example, a central hypothesis is that human operators increase their proficiency with repeated execution of an action and, in support of this, the authors cite Crossman's finding that performance proficiency in a task continues to increase even after 10,000 trials. This may be fine when performance is defined in terms of skill proficiency, but task performance is affected by many more variables than is the execution of skilled behaviour. No simple statement can be made about task performance (e.g. that it degrades or improves over time) unless one knows something about the performer and the nature of the task. While the authors have been careful to provide evidence for the validity of their work by comparing their group simulation model with the actual performance of a submarine crew on a 21-day training mission, there is still the question of whether the predictive value of the simulation could be improved by using additional, and perhaps even different, social-psychological concepts. But perhaps these psychological criticisms are unnecessarily chauvinistic when set against the book's major aim of showing how digital simulation can be utilized in the study of complex socio-technical systems. In this latter respect, the book is very instructive.

ROBERT COOPER

Ethology and Society: Towards an Anthropological View. By HILARY CALLAN. Oxford: Clarendon Press. 1970. Pp. 176. £2.00.

The author examines the possibility of establishing links between social anthropology and the 'classical' ethology of Lorenz and Tinbergen, so as to bring new ideas and new 'social facts' into a theory of society. While constantly warning against the dangers of circularity she notes that much of the attractiveness of this particular link lies in apparent analogies between social processes in animals and man and the common vocabulary, used by ethologists and social anthropologists, to describe them.

After a critical outline of the antecedents of both disciplines and earlier contacts between them, she goes on to discuss the levels at which such contact is, at the present, possible and likely to be fruitful. With these in mind the main body of the book considers the problem of whether, for

[illegible] the ethological concept of aggression can illuminate the problem of social control in anthropology and similarly whether the ethological concept of [illegible] and work on [illegible] throw any light on the problem of the position of women. There are also chapters on 'grooming' and on regulating the size of social groups.

[illegible] optimistic about the outcome of work on these lines and as she is concerned with mapping out a programme of research one can look forward to some more detailed work later. The book is written mainly for the social anthropologist, and the more ethologically inclined reader may find that the description of animal behaviour in terms derived from everyday human life has already made its contribution and given way to a slightly more objective approach in some areas. This is, however, a carefully written, intelligent and critical book and any ethologists still addicted, as Mrs Callan puts it, to the 'occupational habit ... of throwing out vague but inspired remarks about the significance of their findings in the human context' might do well to read it carefully before venturing into print again.

MARGARET VINCE

Br. J. Psychol. (1971), 62, 3, pp. 435–438
Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review.)

Bret, V. J. [illegible] *Learning: from Research to Teaching*. Hemel Hempstead: Prentice-Hall. 1970. Pp. viii + 152. £1.65.

Belliveau, F. & Richter, L. *Understanding Human Sexual Inadequacy*. London: English Universities Press. 1970. Pp. 242. £2.00.

Beyner, K. W. *Wirksamkeit der Gesundheitserziehung*. Basel: S. Karger. 1970. Pp. 70. 19 S.F.

Bühler, C. & Massarik, F. (eds.). *The Course of Human Life*. New York: Springer. 1968. Pp. ix + 422. Price not given.

Chesser, E. *Human Aspects of Sexual Deviation*. London: Jarrold. 1971. Pp. 256. £2.25; paper, 40 np.

Child Development, Society for Research in. *Cognitive Development in Children*. (Five monographs.) London: University of Chicago Press. 1970. (First one-volume edition.) Pp. xii + 698. £4.85.

Côme, D. *Les Obstacles à la germination* (Monogr. de Physiologie Végétale, no. 6.) Paris: Masson. 1970. Pp. 162. 60 FF.

Corah, N. L. & Gale, E. N. *The Origins of Abnormal Behavior*. London: Addison-Wesley. 1971. Pp. xiv + 448. £3.05.

D'Ambrosio, R. *No Language but a Cry*. London: Cassell. 1970. Pp. 252. £2.20.

De Jong, R. N. & Sugar, O. (eds.). *Neurology and Neurosurgery Year-Book 1970*. Chichester: Wiley. 1970. Pp. 512. £6.55.

Deleuze, G. *Sacher-Masoch: An Interpretation, plus Venus in Furs*. London: Faber. 1971. Pp. 248. £2.50.

Eklund, H. *Multiple Choice and Retention*. Uppsala: Almqvist & Wiksell. 1968. Pp. 223. Price not given.

Elonen, A. S. *Performance Scale Patterns in Various Diagnostic Groups*. Jyvaskylä (Finland): Jyväskylä Studies in Education. 1968. Pp. 53. Price not given.

European Committee on Crime Problems. *Collected Studies in Criminological Research*. Vol. v. Strasbourg: Council of Europe. 1970. Pp. 82. Price not given. Vol. vi. Pp. 496. Price not given.

Ewing, A. W. G. *Aphasia in Children*. New York: Hafner. 1967. (Originally published 1930.) Pp. 163. $7.00.

Fieandt, K. v. & Harlahti, J. *New Studies on Perimetric Thresholds of Pigment Colours*. Helsinki: Suomalainen Tiedeakatemia. 1970. Pp. 63. 7.50.

Finkle, R. B. & Jones, W. S. *Assessing Corporate Talent: a Key to Managerial Manpower Planning*. Chichester: Wiley. 1970. Pp. ix + 248. £4.75.

Fitzgerald, H. E. & McKinney, J. P. *Developmental Psychology: Studies in Human Development*. Homewood, Ill.: Dorsey Press. 1970. Pp. xi + 591. $8.95.

Friedman, A. S. *Therapy with Families of Sexually Acting-out Girls*. New York: Springer. 1971. Pp. x + 214. $7.50.

Gathercole, C. E. *Assessment in Clinical Psychology*. Harmondsworth: Penguin Books. 1968. Pp. 174. 30 np.

Gesell, A. *et al.* (eds.). *The First Five Years of Life: a Guide to the Study of the Pre-School Child*. London: Methuen. 1971. (First published 1950.) Pp. 393. £2.00.

G[illegible], P. *La Formation continue des adultes*. Paris: Presses Universitaires [illegible] Pp. [illegible] 13 FF

Gratiot-Alphandéry, H. & Zazzo, R. *Traité de psychologie de l'enfant, 4: [illegible] affectif et moral*. [illegible] Gratiot-Alphandéry.) Paris: Presses Universitaires. [illegible] Pp. [illegible] 18 FF.

Greater London Council. *The G.L.C. on the Environment*. London: G.L.C. [illegible] Pp. 16.

Haas, W. *Phonographic Translation*. Manchester: Manchester University Press. [illegible] Pp. 94. £1.50.

Haas, W. (ed.) *Alphabets for English*. Manchester: Manchester University Press. [illegible] Pp. [illegible] £1.25.

Hajek, J. *Nonparametric Statistics*. San Francisco: Holden-Day. 1969. Pp. [illegible] + 184. $12.00.

Harkavy, O. & Maier, J. *Research in Reproductive Biology and Contraceptive Technology: Present Status and Needs for the Future*. New York: Ford Foundation. [illegible] Pp. 48. Price not given.

Harris, T. A. *The Book of Choice: a Practical Guide to Transactional Analysis*. London: Cape. 1970. (First published 1967.) Pp. xix + 278. £1.75.

Hawton, H. *Controversy: the Humanist Christian Encounter*. London: Pemberton. 1971. Pp. [illegible] £1.25 and 60 np.

Hayrynen, Y. P. *The Flow of New Students to Different University Fields: Career Motivation, Educational Choice and Discriminating Effects of University Admissions: a Study of Finnish Female Students*. Helsinki: Suomalainen Tiedeakatemia. 1970. Pp. 142. 20.00.

Hürsch, L. *Der Einfluss verschiedener Versuchssituationen auf die Faktorenstruktur von Intelligenzleistungen*. Bern: Huber. 1970. Pp. 151. D.M. 29.

India, Govt. of. *Human Problems of Shift Work*. New Delhi: Shri Ram Centre for Industrial Relations. 1971. Pp. 247. $7.00.

Janis, I. L. *et al.* (eds.) *Personality: Dynamics, Development, Assessment*. New York: Harcourt, Brace & World. 1969. Pp. xxxii + 859. £4.75.

Joseph, N. R. *Physical Chemistry of Aging*. London: Academic Press. 1971. Pp. x + 112. £4.45.

Kaplan, M. J., Colarelli, N. J., Gross, R. B., Leventhal, D. & Siegel, S. M. *The Structural Approach in Psychological Testing*. Oxford: Pergamon Press. 1970. Pp. xi + 195. £3.50.

Kaufmann, H. *Aggression and Altruism: a Psychological Analysis*. New York: Holt, Rinehart & Winston. 1970. Pp. x + 165. £1.20.

Kessel, F. S. *The Role of Syntax in Children's Comprehension from Ages 6 to 12*. Chicago: University of Chicago Press. 1970. Pp. 95. Price not given.

Kessen, W. & Kuhlman, C. (eds.) *Thought in the Young Child*. London: University of Chicago Press. 1970. Pp. 146. £1.35 (paper).

Kiell, N. *The Universal Experience of Adolescence*. London: University of London Press. 1969 (2nd edition.) Pp. 942. £6.25.

Klasen, E. *Das Syndrom der Legasthenie, unter besondere Berucksichtigung physiologischer, psychopathologischer, testpsychologischer und sozialer Korrelate*. Bern: Huber. 1970. Pp. 310. D.M. 29.

Kovac, D. (ed.). *Visuelles Wahrnehmen: sein Anteil am Verhalten*. Bratislava: Slowakischen Akademie der Wissenschaften. 1970. Pp. 181. Price not given.

La Fauci, H. M. & Richter, P. E. *Team Teaching at the College Level*. Oxford: Pergamon Press. 1970. Pp. x + 157. £2.75.

Lenneberg, E. H. *Biological Foundations of Language*. Chichester: Wiley. 1967. Pp. xvi + 489. £6.00.

Liebert, R. M. & Spiegler, M. D. *Personality: an Introduction to Theory and Research*. Homewood, Ill.: Dorsey Press. 1970. Pp. xi + 426. $9.00.

[illegible], H. C. & [illegible], D. *Psychology: an Introduction to a Behavioral Science.* [illegible] edition. Chichester: Wiley. 1971. Pp. xvi + 759. £3.50.

[illegible], H. C., [illegible], D. & [illegible], F. (eds.). *Current Research in Psychology: a Book of [illegible].* Chichester: Wiley. [illegible]. Pp. [illegible]. [illegible] and [illegible].

[illegible], G. & [illegible], E. (eds.). *Handbook of Social Psychology.* 2nd edition. Vol. 2. *Research Methods.* London: Addison-Wesley. 1968. Pp. 819. $12.50.

[illegible], J. & [illegible], D. (eds.). *[illegible] of [illegible] Psychology.* London: Holt Rinehart & Winston. 1968. Pp. x + 644. £4.70.

[illegible], A. J. *Dealing with Data.* Oxford: Pergamon Press. [illegible]. Pp. xvi + 192. £2.50 and £1.75 (flexi).

Marks, I. M. *Fears and Phobias.* London: Heinemann. 1969. Pp. 302. £2.25.

M[illegible], I. *[illegible] Dynamics of Learning.* Stockholm: Almqvist & Wiksell. 1970. Pp. 109. Kr. 34.

Maxwell, A. E. *Basic Statistics in Behavioural Research.* Harmondsworth: Penguin Books. 1970. Pp. 126. 35 np.

Medinnus, G. R. & Johnson, R. C. *Child and Adolescent Psychology: a Book of Readings.* Chichester: Wiley. 1970. Pp. ix + 662. £5.00 and £2.65.

Meredith, P. *Parents, Politics, and Protests: a Black and White Paper on Education.* London: University of Leeds. 1970. Pp. 44. 25 np.

Ord, I. G. *Mental Tests for Pre-Literates.* London: Ginn. 1971. Pp. xiii + 270. £2.75.

Passwater, R. A. *Guide to Fluorescence Literature.* Vol. 2. London: Plenum. 1970. Pp. 369. $22.50.

Ramsey, P. H. (ed.). *The Price Revolution in 16th Century England.* London: Methuen. 1971. Pp. ix + 182. £2.00 and 60 np.

Rapaport, D., Gill, M. M. & Schafer, R. (Ed. by R. Holt.) *Diagnostic Psychological Testing.* (Revised edition.) London: University of London Press. 1970. Pp. x + 562. £6.50.

Rollin, P. *Phytochrome: Photomorphogenèse et photopériodisme.* Paris: Masson. 1970. Pp. 136. 60 FF.

Rycroft, C. *Reich.* London: Fontana. 1971. Pp. 104. 30 np.

Sanford, F. H. & Wrightsman, L. S., Jr. *Psychology: a Scientific Study of Man.* (3rd edition.) Belmont, Calif.: Brooks/Cole. 1970. Pp. 736. Price not given.

Scrimshaw, N. S. & Gordon, J. E. (eds.). *Malnutrition, Learning and Behavior.* Cambridge, Mass.: M.I.T. Press. 1969. Pp. 566. £5.85.

Shields, R. W. *A Cure of Delinquents: the Treatment of Maladjustment.* (2nd edition.) London: Heinemann. 1971. Pp. 191. £1.50.

Small, L. *The Briefer Psychotherapies.* New York: Brunner/Mazel. 1971. Pp. xx + 262. $7.50.

Spuhler, J. N. *Genetic Diversity and Human Behavior.* Chicago: Aldine. 1968. Pp. 301. $7.50.

Tanner, J. M. & Inhelder, B. (eds.). *Discussions on Child Development.* Vols. I–IV. London: Tavistock Publications. 1971. (One-volume edition.) Pp. 600. £2.50.

Thompson, H. *Seeker at the Gate.* London: Churches Fellowship for Psychical and Spiritual Studies. 1971. Pp. 125. 90 np.

Vergote, A. *The Religious Man: a Psychological Study of Religious Attitudes.* Dublin: Gill & Macmillan. 1969. (First published as *Psychologie religieuse* in Brussels.) Pp. vii + 306. £3.00.

Vernon, P. E. *The Structure of Human Abilities.* London: Methuen. 1971. (First published 1950.) Pp. 208. 90 np.

Walker, E. L. *Psychology as a Natural and Social Science.* Belmont, Calif.: Brooks/Cole. 1970. Pp. 145. Price not given.

WECHSLER, D. *The Range of Human Capacities.* (2nd edition, revised.) Darien, Conn.: Hafner 1969. Pp. ix + 190. $9.50.

WEINSCHENK, C. *Rechenstorungen · ihre Diagnostik und Therapie.* Bern: Huber. 1970. Pp. 151. D.M. 24.

WILLIAMS, N. & S. *The Moral Development of Children.* Basingstoke: Macmillan. 1970. Pp. 122 60 np.

WOLLHEIM, R. *Freud.* London: Fontana. 1971. Pp. 240. 40 np.

CHANGE-OVER TO SI UNITS

From January 1972 the Journals of the British Psychological Society will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Further information about SI units is contained in the revised *Suggestions to Authors* pamphlet issued by the Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

Authors who refer to physical measures in their paper should now normally use SI units; common units of time (e.g. hour, year) will, of course, persist. Conversion tables will be published during the change-over period, but in the *British Journal of Psychology* these will be restricted to units of length, as others rarely occur. When they do, conversion factors will be stated in footnotes.


AC POWER PACK MODEL 9502
VRB STIMULUS UNIT

Journal of the Experimental Analysis of Behavior

The *Journal of the Experimental Analysis of Behavior* is primarily for the original publication of experiments relevant to the behavior of individual organisms. Review articles and theoretical papers will also be considered for publication. Invited book reviews appear on occasion.

During 1970 over 800 pages of experimental studies were published in JEAB's bi-monthly issues. Below are a few of the articles of special interest to readers of THE BRITISH JOURNAL OF PSYCHOLOGY

Bastian, Chris W. and Hothersall, David. Operant conditioning in redwinged blackbirds.

Burks, Charles D. Schedule-induced polydipsia: are response-dependent schedules a limiting condition?

Delini-Stula, Alexandra. The development and the extinction of hyperthermia induced by conditioned avoidance behavior in rats.

Dews, P. B. Drugs in psychology: a commentary on Travis Thompson and Charles R. Schuster's *Behavioral Pharmacology*.

Fay, Richard R. Auditory frequency generalization in the goldfish (*Carassius auratus*).

Forgione, Albert G. The elimination of interfering response patterns in lever-press avoidance situations.

Goldberg, Steven R. and Schuster, Charles R. Conditioned nalorphine-induced abstinence changes: persistence in post morphine-dependent monkeys.

Herrnstein, R. J. On the law of effect.

Irwin, R. J. and Terman, Michael. Detection of brief tones in noise by rats.

Lyon, David O. and Ozolins, Delmar. Pavlovian conditioning of shock-elicited aggression: a discrimination procedure.

Myers, William A. Observational learning in monkeys.

Pierrel, Rosemary; Sherman, J. Gilmour; Blue, Scot and Hegge, Frederick W. Auditory discrimination: a three-variable analysis of intensity effects.

Premack, David. A functional analysis of language.

Rosenblith, Joan Z. Polydipsia induced in the rat by a second-order schedule.

Siegel, R. K. Apparent movement detection in the pigeon.

Terman, Michael. Discrimination of auditory intensities by rats.

Tracy, William K. Wavelength generalization and preference in monochromatically reared ducklings.

Subscriptions

$16.00 institutions
8.00 individuals (personal use only)
4.00 students

No additional postage is required. Students and other individuals wishing to take advantage of our special rates should enter their subscriptions directly with the Business Office below.

Address subscription orders to

Ms. Kay Dinsmoor, JEAB
Department of Psychology
Indiana University
Bloomington, Indiana 47401 USA

REVUE DE PSYCHOLOGIE APPLIQUÉE

Publication Trimestrielle

Directeurs: Dr P. PICHOT et P. RENNES

1er Trimestre 1971 – Volume 21 – No. 1

SOMMAIRE

2ème Trimestre 1971 – Volume 21 – No. 2

SOMMAIRE

The British Journal of Psychology

The British Journal of Medical Psychology

The British Journal of Social and Clinical Psychology

The British Journal of Mathematical and Statistical Psychology

The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology, *The British Journal of Medical Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts, the subscription price, payable in advance, being £7.50 per volume post free ($25.000 in the U.S.A.) for *The British Journal of Psychology*, and £6.50 each ($22.00) for *The British Journal of Medical Psychology* and *The British Journal of Social and Clinical Psychology*. Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £6 per volume post free ($18 in the U.S.A.). Subscriptions should be sent to the Society's Agents, John Wright & Sons Ltd., 42–44 Triangle West, Bristol BS8 1EX; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year in February, June and November. The subscription is £4 a year post free, payable in advance, or £1.50 each number. The subscription price in the U.S.A. is $8.40. Subscriptions should be sent to the Subscription Dept Manager, Associated Book Publishers Ltd, Distribution Centre, North Way, Andover, Hants.

The series of *Monograph Supplements* issued in connexion with *The British Journal of Psychology* is now discontinued. Orders may be sent to any bookseller, or Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor W. Sluckin, Department of Psychology, The University, Leicester, LE1 7RH;

for *The British Journal of Medical Psychology* should be sent to

Dr J. J. Sandler, Room 101, Department of Psychiatry, Institute of Psychiatry, De Crespigny Park, London, S.E. 5;

for *The British Journal of Social and Clinical Psychology* should be sent to

Professor Thelma Veness (Social Psychology Editor), University of London Institute of Education, Malet Street, London, W.C. 1, *or* to
Dr J. G. Ingham (Clinical Psychology Editor), M.R.C. External Staff, Llandough Hospital, Penarth, Glamorgan, CF6 1XW.

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Dr A. R. Jonckheere, Department of Psychology, University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Professor J. D. Nisbet, Department of Education, University of Aberdeen, Old Aberdeen.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £8.50. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1.50 annually), *The British Journal of Medical Psychology* (£1.50 annually), *The British Journal of Social and Clinical Psychology* (£1.50 annually), *The British Journal of Mathematical and Statistical Psychology* (£1.50 annually), and *The British Journal of Educational Psychology* (£1.50 annually).
Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18–19 Albemarle Street, London, W1X 4DN.

CONTENTS

Printed in Great Britain at the University Printing House, Cambridge

Volume 62 Part 4 [illegible]

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by W. SLUCKIN

with J. ANNETT, W. S. ANTHONY, J. BROWN, R. CONRAD, N. F. DIXON, D. H. FRASER, J. A. GRAY, I. MARTIN, R. L. REID, H. R. SCHAFFER *and* P. B. WARR

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB

AMERICAN BRANCH:

32 EAST 57TH STREET, NEW YORK, N.Y. 10022

£2.50 net $8.50

Preparation of Papers for The British Journal of Psychology

In submitting papers for publication in *The British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. **56** (1965) onwards.

A brief summary, suitable for use as an abstract, should precede the article.

Manuscripts should be submitted in duplicate.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from **diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.**

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation: and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.

Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing [illegible] over the diagram. D[illegible] **reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.**

The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* 26, 231–248.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *Année psychol.* 50, 83–96.

World list abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

Subscriptions

***The British Journal of Psychology* is published quarterly.**

Annual subscription (including postage) £7.50 net ($25.00 in the U.S.A.).

Second class postage paid at New York, N.Y.

Published by Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB and 32 East 57th Street, New York, N.Y. 10022.

Br. J. Psychol. (1971), 62, 4, pp. 439–448
Printed in Great Britain

LONG-TERM RETENTION AND TRANSFER OF AN INDUSTRIAL SEARCH SKILL

By K. D. DUNCAN
Department of Psychology, University of Hull

[illegible]

In the late 1930s Bunch and his associates reported a series of experiments demonstrating that transfer effects remained fairly constant over periods as long as 90 days, whilst retention of the learning task declined (Bunch, 1936; Bunch & McCraven, 1938; Bunch & Lang, 1939). Since this work, retention and transfer over long periods and possible explanations of this striking difference have received little attention. In their review, McGeoch & Irion (1952) stressed the need for investigations of the effect 'over a considerably wider range of learning situations'. In the event, subsequent research has added little to the learning tasks originally studied by Bunch and his colleagues. This paper reports a 6-month comparison of retention and transfer of training in the solution of search problems.

In his original study, Bunch used Peterson's 'rational learning' problem (Peterson, 1918), in which a range of numbers is specified and the subject must learn to pair each number in the range with a different letter in a set of letters successively presented to him. Transfer to a second problem remained roughly constant over intervals of 2, 14, 30 and 90 days; e.g. error savings varied slightly around 40 per cent, whereas error savings in relearning the original problem declined to approximately 90 per cent after 30 days and to 60 per cent after 90 days (Bunch, 1936). Essentially similar findings were reported for paired-associate learning of nonsense syllables in an A–B, C–D transfer experiment (Bunch & McCraven, 1938). More recent work, using Gibson's (1941) visual patterns as stimuli, has shown a similar effect over 28 days with an A–B, C–B paradigm, provided that the response is either meaningful (on Noble's 1952 scales) or, if a nonsense syllable, one which has been rendered familiar by pretraining (Ellis & Burnstein, 1960; Ellis & Hunter, 1960, 1961). Gladys (1960), working with children, also found no significant differences in amount of A–B, C–B

transfer over 14 days using single-syllable, four-letter nouns from school readers. Both Bunch (1941) and Ellis (1965) conclude that learning which is specific to a task is the relatively perishable component.

Data consistent with this conclusion were obtained in a recent investigation of an industrial fault-location task in which approximately half the subjects were tested on a transfer criterion and the remainder on a retention criterion (Duncan, 1969). However, besides no significant decline in transfer, no significant decline in retention was found in some of the subjects. This was true of a group who had practised the problem without any guidance about the efficient solution, but not of subjects who had practised with the aid of a decision tree (similar to that illustrated in Fig. 1), which effectively reduced the problem to following a specific procedure. In these subjects there was significant evidence of forgetting. This suggests that subjects can acquire rather stable techniques of problem-solving and that stability of training effects over time does not depend on whether the criterion is the original or a transfer task, but rather on the availability of specific information during practice. Subjects who practised with the decision tree may have learned the specific search routine at the expense of search strategy, in which case it would seem that learning a search routine is relatively poorly remembered, whereas a search strategy, once acquired, is rather stable over time. However, these are *post hoc* interpretations of an experiment which was designed primarily to measure the effects of different training regimes on retention and transfer, rather than decrements over time. The number of subjects for each combination of time interval and training regime was small.

The present investigation attempts a more adequate demonstration of whether or not the difference between retention and transfer previously found in simpler tasks is also found after learning a complex search task. To this end, a single training condition was employed and independent groups of subjects tested after different time intervals. The hypothesis that search strategy is relatively well remembered, compared with specific search routines, is tested indirectly. All subjects practised search problems with a decision tree and were encouraged to learn search strategy as well as the efficient search routine for the particular problem. Thus subjects may learn either search strategy or the specific routine, or both. Ability to learn search strategy may be expected to vary. No assumptions are made about the nature of this ability except that it may be operationally defined by initial unaided performance. This ability should largely determine the amount of transfer and, if search strategy is well remembered, should do so regardless of the interval between training and transfer testing. The amount of retention, on the other hand, may be determined by remembering either the specific decision tree routine or search strategy, or both. The decline in retention over time should therefore be less marked to the extent that the subject is able to learn search strategy. To the extent that it is determined by remembering the specific decision-tree routine, retention should decline. It was therefore predicted that: (1) *retention* would decline over time, whereas *transfer* would not; (2) ability, operationally defined by initial unaided performance, should correlate more highly with *transfer* than with *retention*; (3) the greater the subject's ability to learn search strategy, defined in this way, the less *retention* would decline.

Method

The search problem. The search problem, illustrated in Fig. 1, consists of finding which of a relatively small number of units is 'faulty'. Like the problem described by Detambel & Stolurow, it incorporates both differing probabilities of failure and differing costs of 'tests' (Detambel, 1956; Detambel & Stolurow, 1957). It differs in these respects from the search problems used by Goldbeck and his associates (1957) and more recently by Cohen & Meudell (1968), which all involved rather large numbers of initially equiprobable sites of the object to be located. The problem is also *structured*, like some of the problems investigated by Dale (1958), i.e. units upstream of a 'good' test may be presumed to be good.

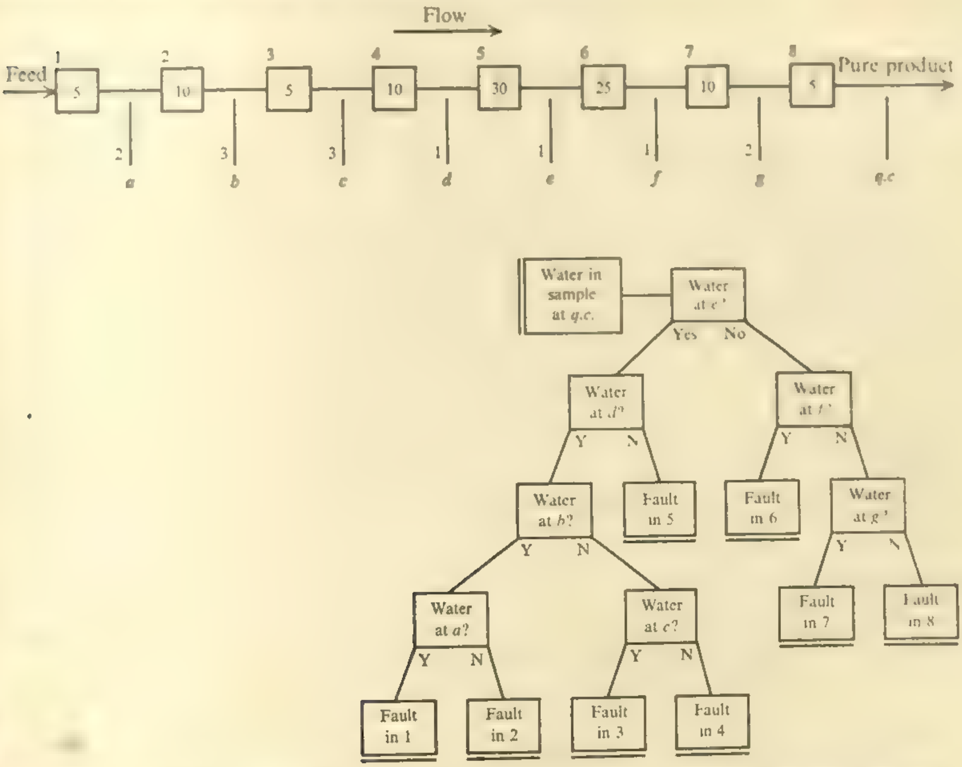


Fig. 1. The diagram illustrates one feature of an industrial acid-purification process. The acid passes through a series of heat exchangers, boxes 1–8, any one of which may leak. The consequent water contamination is initially detected at the quality-control sampling point, *q.c.* A leaking heat exchanger is isolated by taking further samples at points *a*–*g*; thus a sample at *d*, for instance, which is not contaminated, eliminates boxes 1–4, leaving boxes 5–8 still suspect. Because cost of sampling varies in chemical plant, each sampling point has a corresponding cost index of 1, 2 or 3. The figure inside each box is the approximate failure probability. The decision-tree minimizes sampling costs over a random series of faults occurring with frequencies corresponding to their specified probabilities.

Subjects. Subjects were army apprentices, age 16–17. Of the 78 available, 18 were assigned at random to the control, testing only, condition. The remainder underwent training followed by testing. Each subject was assured that the task he was to perform was not a test which would be taken into account at any stage of his career, that he was taking part in a study of industrial skill and that his performance would be treated as confidential and not communicated to any of his military superiors.

Training. The subject was shown a diagram of the series of heat exchangers in Fig. 1 (ex-

cluding the decision tree. In addition to probabilities and costs, the diagram showed the various types of heat exchanger, e.g. 'evaporator reboiler', 'preheater', and of each sampling point, e.g. 'pump drain valve', 'waste heat boiler drain valve'. The direction of flow in the units in the purification process and the quality control (q.c.) sampling point were indicated. It was explained that the presence of water in the q.c. sample indicated a leak in one of the heat exchangers. The subject was then instructed as follows: 'The figures in the boxes, 5, 15, [illegible] out the figures and touching each one with a pencil] these are the chances in 100 of a fault occurring. In other words, suppose over a long period there were altogether 100 incidents of water contamination; about five of them would be traced to a leak in the first unit, the evaporator reboiler [pointing to this unit], 25 would be traced to a leak in number 6, the waste heat boiler, and so on. All right, if you were looking for a fault, what are the chances that it would be in number 7? [If necessary correcting the subject by saying '10' and pointing to the figure in box 7.] Right, what are the chances that it would be in number 3? [again correcting the subject if necessary]. Good. Now when the laboratory reports water in the product line [pointing to the *q.c.* sampling point], the operator finds how it is getting in by taking samples and sending them to the lab. for testing. For instance, if he sent a sample from *c* and the report was 'O.K., no water present', he would know that the fault was not in one of these units [pointing to units 1–3] but in one of these [pointing to units 4–8]. Or if he sent a sample from *g* and the report was water present, he would know that the fault was in one of these units [pointing to units 1–7] and not in this one [pointing to box number 8]. So by taking samples he finds where water is getting in, but he must also take account of the *cost* of the sampling. A sample at *g* costs 2 because it is moderately expensive to sample there. What does a sample at *f* cost? [correcting the subject as before, if necessary]. So the larger the figure at the sampling point, the greater the cost of taking a sample.' The subject's attention was then drawn to two lists of faults, on different coloured paper, lying face down on the table and he was invited to indicate which one should be used first for practice in fault location. He was told that the reason for asking him to choose would be explained later. He was then instructed as follows: 'Now, to trace faults you ask if there is water at *a*, *b*, *c*, *d*, *e*, *f* or *g* and I will say yes or no. The faults on my list do not occur in any special order, so you cannot guess what is coming next, but how *often* each comes up corresponds roughly to the chances shown in each of the boxes [pointing to each unit]. Remember the larger the figure for one of the units, the greater the chances of failure and the more often it is likely to come up on my list. Also bear in mind the different costs of the sampling points [pointing to the cost index at each sampling point]. I will time you and record the samples you take. The time you take is *not* important. *The total cost of the samples you take in finding these faults is what matters*—the smaller the total cost, the better, of course. All right then, the lab. report water at the end of the line here [pointing to the *q.c.* sampling point]. Ask: is there water at *a*, *b*, *c*, *d*, *e*, *f* or *g*, wherever you like, and I will tell you yes or no.'

After the subject had located 10 faults, he was shown the list he had chosen on which the samples he had asked for were now recorded opposite each fault. He was then given a diagram of the decision tree in Fig. 1, shown how it determined the sequence in which samples should be taken and asked to use it to locate the faults on the remaining list. He was then told: 'The guide is an ideal solution to this search problem. You will almost certainly find that it reduces your costs of sampling in finding the next set of faults. After you have used it several times you will probably find that you can remember the sampling order without looking at the guide. Try to do this. Later some of you will be asked to find faults in a similar set of chemical plant units, but where the costs of sampling and the chances of a fault occurring are different. So try to discover *why* taking samples in the way the guide does reduces overall costs; why is it efficient to start at *e*, then go to either *d* or *f*? If the numbers in the boxes and at the sampling points were different, this guide would not be very efficient, but it would always be possible to work out some order of sampling which would be efficient.' The subject was invited to study the guide for 2 min., then the instructions for use of the guide were repeated.

When the subject had located a further 10 faults following the prescribed sampling sequence, he was shown the samples he had taken on the second list and the aggregate costs and was invited to compare the aggregate costs for the two lists. It was explained that the choice of which list to use for unaided practice and which for practice with the guide had been left to him in order to avoid any suspicion of 'rigging' the series of faults to favour the guide. The guide was then removed and the subject located further faults following the prescribed sequence from memory. Departures from the sequence were corrected and training continued until the subject

[illegible]

[illegible] *a* to *g* were 1, 1, 3, 3, 2, 1, 2 respectively.

[illegible] random order, each fault occurring with a frequency corresponding to its specified probability, e.g. faults with five chances in 100 occurred once, faults with 30 chances in 100 occurred six times, etc.

The 18 control subjects were tested in the same way before the training period, nine being assigned at random to the original task and nine to the transfer task.

Scoring errors. In their first attempts to find faults, a few subjects made errors of sampling and errors of location. Errors of sampling were either redundant or repeated samples. If the subject requested a redundant sample he was shown why he did not need it, e.g. the sample previously taken at *d* was contaminated, therefore a sample at *f* or other points downstream of *d* will also be contaminated. If he asked for a repeated sample the subject was simply reminded of his previous finding. If the subject located a fault, correctly or incorrectly, without having taken samples isolating the unit in question, or if he obtained the necessary samples to locate the fault, but nevertheless located it incorrectly, it was explained why his conclusion did not follow from the samples he had taken. If a subject guessed the location of a fault, i.e. before he had isolated it, he was not told whether or not he was correct, but was required to continue until he had obtained the isolating samples.

These errors did not recur in the retention and transfer testing in the 60 subjects who had been trained. However, the control subjects, who were tested on the two tasks without prior training, made some errors in locating the first few faults. Costs of redundant and repeated samples were not added to control group scores. Thus the base-line test scores, against which training effects are assessed, are not enhanced by control group subjects' initial misunderstandings of instructions.

Results

Fig. 2 shows retention and transfer after three intervals between training and testing and also the performance of control groups on the original and transfer tasks, all in terms of excess sampling costs in locating 20 faults. Excess sampling costs is a subject's aggregate costs less a constant of 79 in the case of the retention test, and 78 in the case of the transfer test, these constants being the aggregate costs entailed in locating 20 faults by the most efficient strategy.

Separate unweighted means analyses of variance for the retention and transfer data are shown in Tables 1 and 2. The differences between transfer on the three occasions are small and not significant, although there is some decrement after 58 and 182 days. The differences between retention on the three occasions are significant but only at the 5 per cent level. Most of the variance for this main effect is shown by trend analysis (Winer, 1962) to be attributable to linear regression when appropriate orthogonal polynomials are constructed for the different retention intervals (Robson,

1959). All retention and transfer means are significantly different from the controls using Dunnett's (1955) multiple comparison procedure. Thus after 182 days the effects of training are still present in both retention and transfer data.

There are highly significant differences in retention and transfer between subjects of high and subjects of low ability. Ability to learn search strategy is measured, for this purpose, by the aggregate cost of locating the first 10 faults in which neither

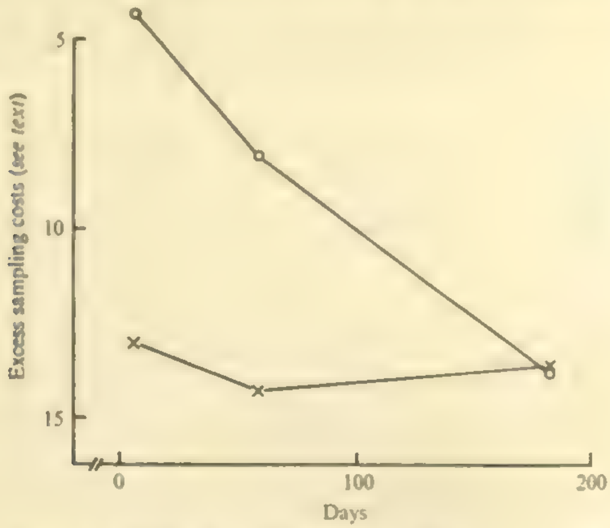


Fig. 2. ○——○, original task (control mean = 24·7); ×——×, transfer task (control mean = 26·8).

Table 1. *Analysis of variance for retention of original task and control group data*

Source	D.F.	M.S.	*F*	*P*
A. Occasions – 6, 58 and 182 days	2	215·184	4·08	< 0·05
B. Ability – high *v.* low	1	1,172·724	22·22	< 0·005
C. Training *v.* no training	1	1,793·913	33·99	< 0·005
AB	2	224·222	4·25	< 0·025
BC	1	236·777	4·49	< 0·05
Simple effects of A				
High ability	2	47·436	0·90	—
Low ability	2	391·971	7·43	< 0·005
Simple effects of C				
High ability	1	363·611	6·89	< 0·025
Low ability	1	1667·079	31·58	< 0·005
Trends in A main effect				
Linear	1	424·505	8·04	< 0·01
Quadratic	1	5·864	0·11	—
Differences in trend for simple effects of A				
Linear	1	223·827	4·24	< 0·05
Quadratic	1	224·613	4·26	< 0·05
Trends in simple effects of A – low ability				
Linear	1	632·417	11·98	< 0·005
Quadratic	1	151·525	2·87	—
Error within cell	30	52·785	—	—

Table 2. Analysis of variance for transfer task and control group data

Source	D.F.	M.S.	F	P
A. Occasions – 6, 58 and 182 days	2	3·996	0·07	—
B. Ability – high v. low	1	3,247·362	54·61	< 0·005
C. Training v. no training	1	1,220·635	20·85	< 0·005
AB	2	17·393	0·29	—
BC	1	36·800	0·63	—
Error, within cell	30	58·543	—	—

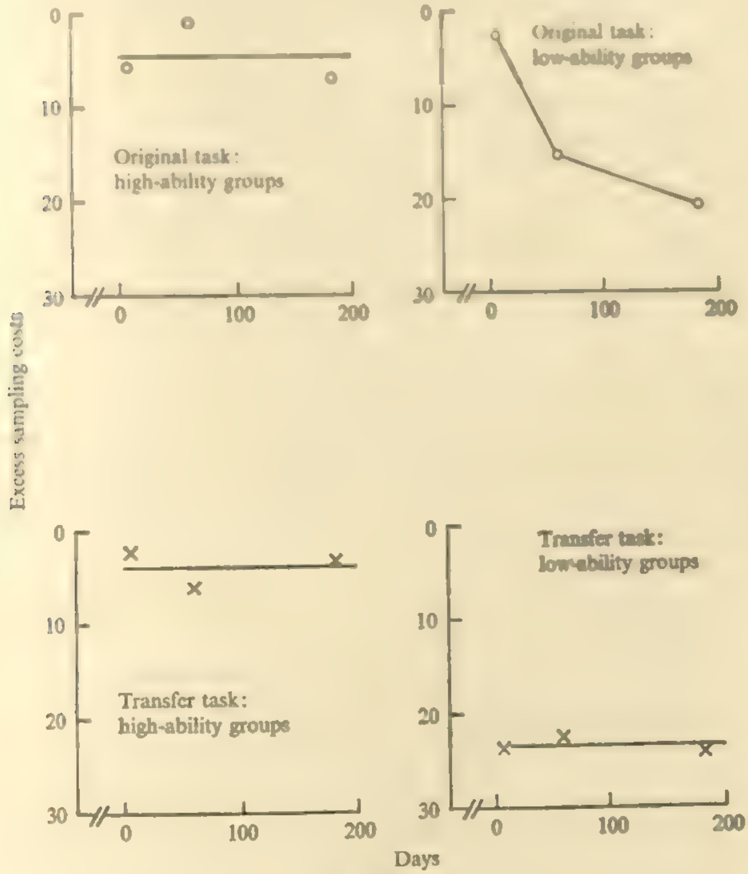


Fig. 3

experimental nor control group subjects were helped in any way. In the retention data there is an interaction between ability and occasions which, in view of the propositions under test, is accounted for in terms of the simple effects of each ability level. The retention of high-ability subjects does not differ significantly over the three occasions, whereas retention of low-ability subjects does, 80 per cent of the variance being attributable to highly significant linear regression on the retention intervals. This is consistent with the proposition that decline in retention would result from subjects forgetting the decision-tree sequence and that this decline would

be less marked in higher-ability subjects because of their superior learning of search strategy. Fig. 3 shows retention and transfer subgroup means at each interval in relation to the trends indicated by analysis of variance.

Similar analyses of variance of mean location time for 20 faults found highly significant differences, in both the retention and transfer tests, between high- and low-ability subjects and between experimental groups taken together and the control; no other effects were significant.

The correlations of ability with retention and with transfer in Table 3 are consistent with predictions and, again, with the basic proposition that search strategy is comparatively well remembered, whereas memory of specific sampling routines is comparatively perishable.

Table 3. *Correlations of ability† with retention and transfer*

Correlation of ability† with sampling costs in locating 20 faults	All groups ($n = 29$)	After 6 days ($n = 10$)	After 58 days ($n = 10$)	After 182 days ($n = 9$)
Original task	0·470*	0·048	0·551	0·777*
Transfer task	0·854**	0·870**	0·880**	0·858**

* $P < 0{\cdot}05$. ** $P < 0{\cdot}01$. † Initial unaided fault location; see text.

DISCUSSION

Transfer of search skill seems to be remarkably stable over a 6-month period. This is also true of retention in abler subjects. Lack of decrement probably depends not so much on whether retention or transfer is the criterion as on the extent to which subjects learn search strategy rather than a specific search sequence. This interpretation is consistent with the conclusion drawn by McGeoch & Irion (1952) from studies of less complex tasks, namely that effects of learning which are stable over time consist of 'general factors, such as modes of attack which are presumably more resistant to forgetting than are specific items'. But the question remains *why* a more general skill should be better retained than specific response sequences.

One possible explanation, in the case of this task, lies in what is assumed to be learned. At one extreme, the subject might only learn the subchains of the decision tree *per se*, in which case the amount of practice received by each subchain can be specified (Table 4). At the other extreme, a subject might learn more general rules of which these chains are exemplars. One rule a subject might learn is to sample such that the sum of fault probabilities on each side of the sampling point are as nearly equal as possible ('half-split'). Another rule a subject might learn is to select cheaper rather than dearer samples. A subject who applied the half-split rule from the occasion on which he first used the decision tree would practise it $20+12+8+6 = 46$ times in the course of 20 search problems (since the four chains in Table 4 are all exemplars of this rule). Similarly, preferring cheaper samples would be practised 40 times (subchains C 1, C 2 and C 3). Clearly individual subchains can be practised on fewer occasions than rules. Subjects may, of course, learn a mixture of both general rules and the specific subchains of the decision tree. The abler the subject (as operationally defined in this paper) the earlier in the training series he would begin to practise rules. Whether what is practised is general or specific, it seems reasonable

to assume that an empirical law of exercise will apply. Rules may be better retained than subroutines of a procedure simply because applying rules receives much more practice and may, indeed, be overlearned.

Whilst not denying the effects of interference or facilitation from antecedent or subsequent learning, it is argued that the intervening learning itself should not be neglected in the interpretation of retention and transfer phenomena. Other studies which have demonstrated stability of transfer over time can also be interpreted in terms of what the subject may have learned. Presumably a subject learning Bunch's (1936) task must continually review which of the set of numbers have not been assigned to previous letters in the series and which of the remainder have been rejected on the current trial. Reviewing and selecting from the set of alternative responses would be practised at least once on every trial. Since Bunch used 10 letter problems, such a process of response selection and the amount of practice it received could reasonably account for the stable transfer effect which he demonstrated.

Table 4. *Frequency of practice of subchains of the decision tree during location of 20 faults**

Subchains† of decision tree in Fig. 1	Times practised
C 1. Water in sample at *g.c.* – water at *e*?	20
C 2. Water at *e*? – yes – water at *d*?	12
C 3. Water at *e*? – no – water at *f*?	8
C 4. Water at *d*? – yes – water at *b*?	6

* Occurring with frequencies proportional to specified probabilities.

† The remaining eleven subchains of the decision tree are specified by the nature of the task as explained in the instructions to subjects.

Ellis & Burnstein (1960) used a two-process model to explain their A–B, C–B data, i.e. response differentiation or integration, followed by associative 'hook-up'. They hold that the associative hook-up component of transfer is stable over time, whereas response differentiation is not – a view which also accounts for the stability over time of A–B, C–D transfer since it has no response differentiation component. Now associative hook-up may be supposed to include some form of scanning and choosing from a limited set of links or linking mechanisms. Such a process would be practised on every trial and might well be stable over time simply because of the extent to which it is practised.

Thus, in other experimental situations where transfer has been found stable over time, it is possible to argue that considerable overlearning occurs – overlearning, that is, of a process which is rather general, in the sense that it is practised more often during learning than individual responses. In the experiment reported here, opportunities to practise search strategy and opportunities to practise the response sequences of a specific search routine can be estimated fairly directly. The extent to which remembering search strategy may be favoured, if it is practised during learning, is considerable. It therefore seems that a simple law of exercise is an economical explanation of the data.

I am indebted to B.P. Chemicals (U.K.) Ltd. for providing facilities to study [illegible] tasks, to Colonel J. W. Eagle, M.B.E., E.R.D., members of the staff and apprentices of the Army Apprentice College, Harrogate, to the Department of Employment and Productivity for financial support, and to Dr H. C. A. Dale for criticizing a draft of this paper.

REFERENCES

BUNCH, M. E. (1936). The amount of transfer in rational learning as a function of time. *J. comp. Psychol.* **22**, 325–337.

BUNCH, M. E. (1941). A comparison of retention and transfer of training from similar material after relatively long intervals of time. *J. comp. Psychol.* **32**, 217–231.

BUNCH, M. E. & LANG, E. S. (1939). The amount of transfer of training from partial learning after varying intervals of time. *J. comp. Psychol.* **27**, 449–459.

BUNCH, M. E. & MCCRAVEN, V. G. (1938). The temporal course of transfer in the learning of memory material. *J. comp. Psychol.* **25**, 487–496.

COHEN, J. & MEUDELL, P. (1968). Strategies of search. *Acta psychol.* **28**, 322–343.

DALE, H. C. A. (1958). Fault finding in electronic equipment. *Ergonomics* **1**, 356.

DETAMBEL, M. H. (1956). Probabilities of success and amounts of work in a multichoice situation. *J. exp. Psychol.* **51**, 41–44.

DETAMBEL, M. H. & STOLUROW, L. M. (1957). Probability and work as determiners of multichoice behaviour. *J. exp. Psychol.* **53**, 73–81.

DUNCAN, K. D. (1969). Task analysis evaluated. In F. Bresson & M. de Montmollin (eds.), *La recherche en enseignement programmé: tendances actuelles*. Paris: Dunod.

DUNNETT, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Am. statist. Ass.* **50**, 1096–1121.

ELLIS, H. C. (1965). *The Transfer of Learning*. New York: Macmillan.

ELLIS, H. C. & BURNSTEIN, D. D. (1960). The effect of stimulus similarity and temporal factors in perceptual transfer of training. *Tech. Rep.* no. 1. Albuquerque, New Mexico: Sandia Corporation.

ELLIS, H. C. & HUNTER, J. E. (1960). Response meaningfulness and the temporal course of transfer. *Tech. Rep.* no. 2. Albuquerque, New Mexico: Sandia Corporation.

ELLIS, H. C. & HUNTER, J. E. (1961). The effect of response familiarization on the temporal course of transfer. *Tech. Rep.* no. 3. Albuquerque, New Mexico: Sandia Corporation.

GIBSON, E. J. (1941). Retroactive inhibition as a function of the degree of generalization between tasks. *J. exp. Psychol.* **28**, 93–115.

GLADYS, M. (1960). Grade differences in transfer as a function of the time interval between learning tasks. *J. educ. Psychol.* **51**, 191–194.

GOLDBECK, R. A., BERNSTEIN, B. B., HILLIX, W. A. & MARX, M. H. (1957). Application of the half-split technique to problem-solving tasks. *J. exp. Psychol.* **53**, 330–338.

MCGEOCH, J. A. & IRION, A. L. (1952). *The Psychology of Human Learning*. New York: Longmans, Green.

NOBLE, C. E. (1952). An analysis of meaning. *Psychol. Rev.* **59**, 421–430.

PETERSON, J. C. (1918). Experiments in rational learning. *Psychol. Rev.* **25**, 443–467.

ROBSON, D. S. (1959). A simple method for constructing orthogonal polynomials when the independent variable is unequally spaced. *Biometrics* **15**, 187–191.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

(*Manuscript received* 24 *August* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 449–457
Printed in Great Britain

VERBAL MEDIATION EFFECTS IN CROSS-MODAL TRANSFER

By J. GARVILL and B. MOLANDER
Department of Psychology, University of Umeå, Sweden

Cross-modal transfer was investigated within the verbal mediation paradigm A–B, B–C, A–C. In the first stage the subjects learned to associate CVC syllables with three-dimensional ceramic objects which were presented either visually or tactually. In stage III transfer was tested by presenting the objects in the one of the two modalities which was not used during the first stage learning. Verbal mediation was found to be an effective aid in the transfer of information between the two sensory modalities. It was also found that an interference effect could be obtained in stage II, although non-verbal stimuli were used in stage I. Transfer appeared to be larger when training started tactually and was tested visually than vice versa. It is suggested that this asymmetric effect is caused by differential verbal mediation.

Recent reviews on studies in cross-modal transfer (e.g. Gibson, 1969; Bjorkman, 1971; von Wright, 1970) emphasize the following findings: (*a*) cross-modal transfer is easily demonstrated with human beings but extremely difficult and perhaps impossible in animals, (*b*) the capacity of cross-modal transfer seems to be a function of age, and (*c*) the transfer effect is often found to be asymmetric (e.g. transfer from touch to vision is larger than transfer from vision to touch).

In explaining these empirical findings, several authors have suggested that verbalization or verbal mediation may play an important role in cross-modal transfer (Gaydos, 1956; Houck *et al.*, 1965; Ettlinger, 1967; Garvill & Molander, 1968). However, it is also argued that not all forms of cross-modal transfer need to be mediated by language or verbal behaviour (Blank & Bridger, 1964; Blank *et al.*, 1968). Blank *et al.*, for instance, trained young children in a form discrimination task either visually or tactually and tested for transfer in the other modality. They obtained transfer from vision to touch but not in the other direction. Although transfer occurred, the children could not label the shapes used and the authors concluded that language was not a necessary mediator for transfer. The soundness of this conclusion can be questioned, however, since no transfer occurred from touch to vision and since the lack of overt verbal responses by no means excludes the possibility of covert verbal responses. They also found, which is somewhat unusual, that tactual training was very fast compared to the visual training. Nevertheless, the experiments by Blank & Bridger (1964) and Blank *et al.* (1968) seem to be the only attempts so far to investigate directly the role of language and verbal mediation in cross-modal transfer. A more systematic approach to the clarification of the role of verbal mediation in cross-modal transfer is needed however. Even if it could be shown that language or verbal mediation is not necessary for cross-modal transfer to occur, it is of course important to investigate to what extent these processes can affect the magnitude of transfer.

It seems that the usual verbal mediation paradigms (Jenkins, 1963; Kjeldergaard, 1968) lend themselves very naturally to such a systematic investigation, especially as many cross-modal experiments performed with human subjects have been de-

signed as paired-associate (P-A) learning experiments, where usually verbal labels or numbers have been associated with shapes during training stages (Gaydos, 1956; Walk, 1965; Björkman et al., 1965; Garvill & Molander, 1968; Moley, 1965). According to recent reviews on verbal learning and transfer, there is overwhelming evidence of verbal mediation effects in the context of P-A learning studies (Horton & Kjeldergaard, 1961). Although these effects have been shown to occur in a variety of P-A learning situations and with different kinds of verbal materials, there are few studies demonstrating verbal mediation with non-verbal material. In some cases where random shapes have been used as stimuli there has even been a failure to obtain a mediation effect (Barclay, 1961).

The main purpose of the present experiment is then to demonstrate that verbal mediation is a plausible explanation for transfer of information between two sensory modalities. This is done by investigating cross-modal transfer in one of the usual verbal mediation paradigms. A secondary purpose is to provide some evidence that effects usually found with verbal material can be obtained with non-verbal material also.

The paradigm selected for this experiment was the simple chaining paradigm A–B, B–C, A–C, which was compared to the usual control paradigm A–B, D–C, A–C. By training the subjects tactually in the first stage and testing visually in the third stage or vice versa, it was expected that verbal mediation effects would be obtained independently of modality order. According to earlier findings, it was also expected that transfer should be larger if the training started tactually than if it started visually. In other words, an asymmetric transfer effect should be obtained. With verbal material it is often found that the second stage learning is slower for the experimental group compared to the control group, because of interference from first stage backward associations (Jenkins & Foss, 1965; Horton & Kjeldergaard, 1961). Interference should be expected in this case too, if implicit verbal representations of non-verbal stimuli are formed during learning.

METHOD

Design. The design consisted of two experimental groups learning the material in accordance with the chaining paradigm A–B, B–C, A–C and two control groups following the usual control paradigm A–B, D–C, A–C. One of the experimental groups started by learning the first stage tactually and was tested visually in the third stage. (In the following the notation T–V: E will be used.) The other experimental group started training visually and was tested tactually (V–T: E). The two control groups were of course trained and tested in corresponding modalities (T–V: C and V–T: C). A summary of the experimental design is presented in Table 1.

Table 1. *The experimental design*

	Stages		
	I	II	III
Tactual–visual conditions			
Experimental group (T–V:E)	A–B*	B–C	A–C
Control group (T–V:C)	A–B	D–C	A–C
Visual–tactual conditions			
Experimental group (V–T:E)	A–B	B–C	A–C
Control group (V–T:C)	A–B	D–C	A–C

* The paired-associate task in each stage is symbolized by A–B, B–C, etc., where the first letter refers to stimuli and the second letter to responses.

[illegible]

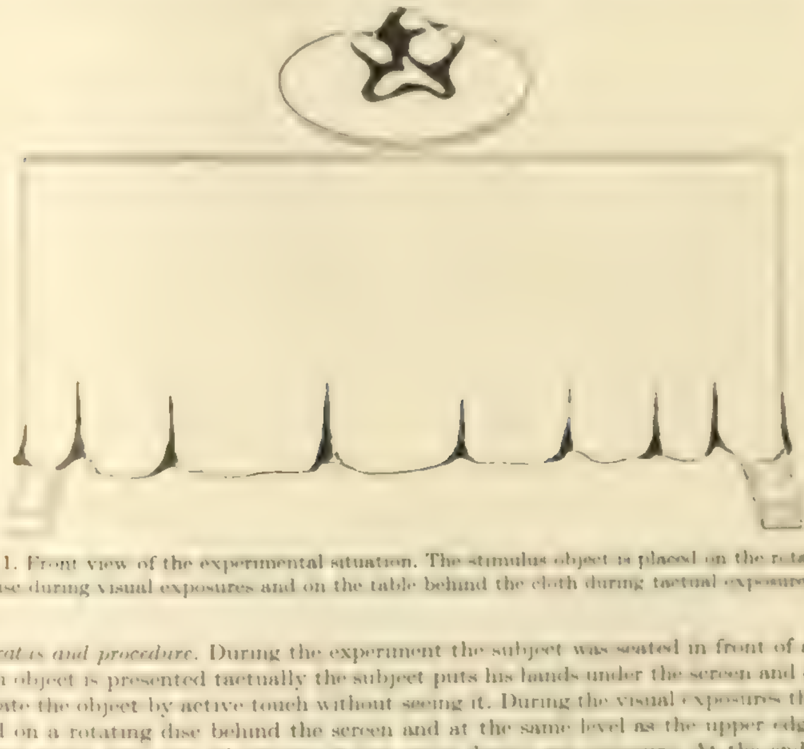


Fig. 1. Front view of the experimental situation. The stimulus object is placed on the rotating disc during visual exposures and on the table behind the cloth during tactual exposures.

Apparatus and procedure. During the experiment the subject was seated in front of a screen. When an object is presented tactually the subject puts his hands under the screen and can then manipulate the object by active touch without seeing it. During the visual exposures the object is placed on a rotating disc behind the screen and at the same level as the upper edge of the screen (eye-level). The rate of rotation is two turns during one exposure. At the end of each exposure which lasts for 10 sec., the assigned syllable is presented by means of a tape-recorder. In order to make it possible for the experimenter to note the answer and select a new object, there was a 10 sec. interval between each exposure. Before the start of first stage learning the experimenter read a standard instruction which explained the anticipation method and pointed out that the learning would go on to a criterion of two successive errorless trials. No information about the following stages was given. A demonstration of the procedure was made by presenting the subject with a stimulus–response pair which did not belong to the experimental pairs. The anticipation method was used in all stages and with the same criterion as in stage I. In order to prevent serial learning, new orders of pairs were employed for every trial.

Before the start of stage II the subject was instructed that there was another task which consisted of pairing syllables and numbers. He was also told to guess the correct response (a number between 1 and 7) for each syllable presented in trial 1. This guessing should occur after the syllable had been presented and before feedback had been given from the tape-recorder. According to anticipation instructions, he was of course expected to attend to the feedback and give as many correct responses as possible in the next and following trials. Each syllable was printed on

a 3 × 5 in. card and shown to the subject for 5 sec. At the end of that interval the correct response was given from the tape recorder. The inter-item interval was 5 sec. In the third stage the task was to pair objects with numbers. The subject was told that the objects would be presented in a different modality. As in stage II, the guessing of responses on trial 1 was employed. The presentation times were the same as in stage I, except that the inter-item interval was 5 sec. The reason for shortening the inter-item intervals in stages II and III was that it was easier for the experimenter to note the answers in these stages. Thus the total time each subject spent in the laboratory could be reduced to about $1\frac{1}{2}$ hr.

Subjects. Forty undergraduate psychology students from the University of Umeå served as subjects. They were randomly assigned to four groups of 10 subjects each. Participation in the experiment was considered as a part of the students' course requirement.

Table 2. *Means and standard deviations of the two criterion measures for each learning stage*

Stage	Group	Number of trials Mean	S.D.	Number of correct responses Mean	S.D.
I	T–V: E	15·3	3·90	4·1	3·14
	T–V: C	12·9	3·11	4·9	3·70
	V–T: E	9·2	3·09	8·7	4·45
	V–T: C	10·7	4·86	8·4	3·72
II	T–V: E	6·4	2·01	17·5	3·53
	T–V: C	5·1	1·64	18·6	4·36
	V–T: E	7·5	2·80	14·8	3·43
	V–T: C	5·2	1·66	18·7	3·61
III	T–V: E	5·3	2·05	21·8	6·23
	T–V: C	8·5	2·46	13·8	2·48
	V–T: E	8·6	3·58	17·8	5·40
	V–T: C	14·4	5·92	11·2	5·00

RESULTS

Two dependent measures were used: number of trials to a criterion of two consecutive errorless trials and number of correct responses during the first four trials. Means and standard deviations for these measures in each of the three stages are presented in Table 2 and Fig. 2.

A three-way analysis of variance with repeated measures on the third factor was performed on each dependent variable. As the analyses yielded the same results for both measures, only results based on the number of correct responses are presented here. The ANOVA was a $2 \times 2 \times 3$ analysis with modality order (I) as the first factor (tactual training and visual test *v.* visual training and tactual test), experimental groups *v.* control groups as the second factor (II), and stage as the third factor (III).

The analysis revealed a significant main effect of the third factor ($F = 122{\cdot}91$; d.f. $= 2, 72$; $P < 0{\cdot}01$), a significant interaction between the first and third factors ($F = 12{\cdot}53$; d.f. $= 2, 72$; $P < 0{\cdot}01$) and a significant interaction between the second and third factors ($F = 22{\cdot}87$; d.f. $= 2, 72$; $P < 0{\cdot}01$). As the main effect is of lesser interest in this context only the interactions were further analysed by applying the Tukey ratio for comparisons among means. This comparison revealed that the I–III interaction (modality order × stage) was due to a significant slower tactual learning compared to the visual learning during the first and third stages ($P < 0{\cdot}05$). This is a common finding when tactual and visual learning is compared on the same material

(Gaydos, 1956, Sinha & Sinha, 1960, Bjorkman *et al.*, 1965, Walk, 1965, Lobb, 1965, Garvill & Molander, 1968, Milne, 1968, Cashdan, 1968). The problem of this differential first stage learning will be discussed later.

The analysis of the II–III interaction revealed that there were no significant differences between experimental and control groups in either modality in the first stage ($P > 0.05$) and no significant difference between control groups and experimental groups in the second stage ($P > 0.05$). The difference in the second stage pointed in the expected direction, however, and was significant at the 10 per cent level. In the third stage the experimental groups were superior to the control groups ($P < 0.01$). The analysis thus indicated a slight interference effect in stage II and a clear mediation effect in stage III. As can be seen from Table 2, T–V: E is superior to

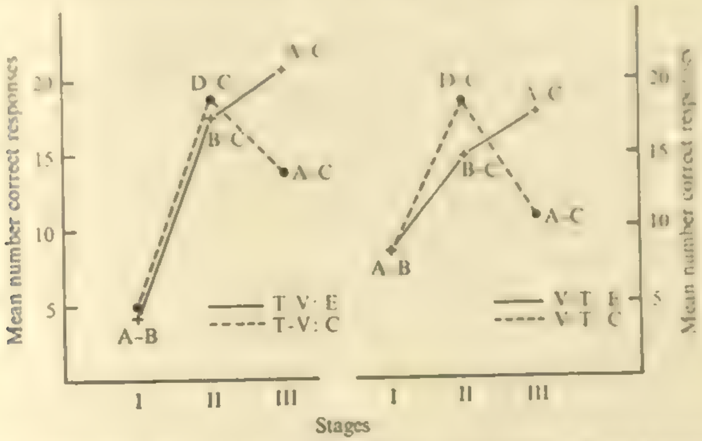


Fig. 2. Mean correct responses in the first four trials for experimental and control groups in each of the three stages. In the left part of the figure the tactual–visual experimental group (T–V: E) is compared with the tactual–visual control group (T–V: C). In the right part of the figure the visual–tactual experimental group (V–T: E) is compared with the visual–tactual control group (V–T: C). A–B, B–C, etc., symbolize the paired-associate tasks in the three stages.

Table 3. *Means and standard deviations of number of correct responses in trial 1 in stages II and III*

	Stage II, trial 1		Stage III, trial 1	
Group	Mean	S.D.	Mean	S.D.
T–V: E	2·0	1·41	4·6	1·91
T–V: C	1·6	1·20	1·0	0·77
V–T: E	1·4	0·80	3·7	1·95
V–T: C	1·6	0·80	1·4	1·02

V–T: E in stage III, which seems to indicate an asymmetric transfer effect. But as was pointed out above, visual learning is faster than tactual learning and a comparison of the two experimental groups would thus involve a confounding of transfer and original learning differences. However, this confounding is minimized on trial 1 in stage III, where the subjects have to guess the correct responses after stimulus presentation but before feedback is given. Data from trial 1 in stage III can be seen in

Table 3. Since there were no differences between the control groups ($P > 0{\cdot}05$), the experimental groups were tested direct by a simple t test. The difference reached significance only at the 10 per cent level, however, thus indicating a slight asymmetric effect.

A possible objection to the asymmetric effect is that it may be a function of different general transfer effects (e.g. warm-up and learning-to-learn). In that case, for instance, the tactual–visual groups have longer learning times in the first stage and thus greater opportunity to acquire an efficient learning strategy. However, in the present design numbers were used as responses in stages II and III, and the subjects are asked to guess in the first trial in these stages. As the numbers 1–7 are used as responses, a complete random guessing should give 1/7 or 0·14 correct responses. On the other hand, subjects obtain as much as 2·59 correct responses ($1/7 + 1/6 + 1/5 + 1/4 + 1/3 + 1/2 + 1/1$) if they attend to the tape-recorder and correctly remember what numbers have been given as feedback. Thus there are indeed possibilities for learning a more efficient guessing strategy. On the assumption of more learning to learn for the tactual–visual groups it would then be expected that T–V: C should have more correct responses in the first trial in stage III compared to V–T: C. But as can be seen in Table 3 there is no difference between these groups ($P > 0{\cdot}05$). Furthermore there is no difference between the performance in trial 1 in stage II and trial 1 in stage III for the same two groups ($P > 0{\cdot}05$). A one-way analysis of variance yielded no difference among the four groups ($P > 0{\cdot}05$) in trial 1 in stage II. These results indicate that the asymmetric effect is not easily explained by differential learning-to-learn. Another possibility is that the asymmetry is caused by differential verbal mediation. It is well known that verbal mediation effects are most pronounced in early trials in the test stage (Jenkins, 1965). Now, if the asymmetry has something to do with verbal mediation, it is expected to be strongest at the beginning of stage III. An indication of the change of mediation effects during the first four trials in stage III can be seen in Table 4, where the differences between experimental and control groups are presented. As can be seen in Table 4, the difference between experimental and control groups decreases during the four trials in both the T–V condition and the V–T condition with a somewhat slower decrease in the latter. It can also be seen that the difference in the T–V condition is larger only during the first two trials, which indicates that the asymmetry decreases when the mediation effect decreases. However, a more thorough investigation of this hypothesis is needed as the effects in Table 4 are rather weak and, except in trial 1, difficult to interpret because of the possibility of confounding.

Table 4. *Mean differences between experimental and control groups during the first four trials in stage III*

	Trial			
	1	2	3	4
T–V: E – T–V: C	3·6	2·6	1·5	0·3
V–T: E – V–T: C	2·3	1·4	1·7	1·2

Discussion

The present experiment gives clear evidence of verbal mediation effects in the transfer of information from one sensory modality to another. The evidence is less clear, however, concerning the expected interference in the second stage and the asymmetry effect in the test stage, although data point in the expected direction. There may be several reasons for these weak effects. One possible source is the difference in rate of learning in stage I. It will be remembered that learning in the tactual modality was significantly slower than in the visual modality. As pointed out by Underwood (1964), there is no way of evaluating the actual association strength of the items at the end of stage I when learning is taken beyond the asymptote. It may well be that the 'true' association strength is stronger among items learned visually because of a faster rate of learning. This statement rests on the assumption that association strength goes on growing after the asymptote has been reached. In view of the fact that overlearning has been shown to improve retention, this seems to be a reasonable assumption. If the A–B associations are weaker in T–V–E compared to V–T–E this would lessen the possibility of interference and mediation in T–V–E as compared to V–T–E. In this context it should be noted that the main part of the interference effect in stage II is concentrated on V–T–E, as can be seen in Fig. 2. According to this reasoning the interference and mediation effects in T–V–E should increase if the association strengths of the two experimental groups could be made equal before the beginning of stage II learning. But in learning experiments where complex forms are used as stimuli this is a very difficult task. On the one hand, there is by now strong evidence that vision has a higher information capacity than touch (Rock & Harris, 1967), which indicates that it is very hard to find a stimulus material which has the same degree of difficulty in the two modalities. If such a material is found it is very likely to be either one- or two-dimensional or very well learned and familiar to subjects pre-experimentally. That kind of material does not, however, as Gibson (1963) has argued, test the limits of the information processing organism, and 'the dimensions of available stimulation in a natural physical environment are of higher order than these, being variables of pattern and change'. On the other hand, the design requires the groups to be equal in first stage learning. Until this dilemma is solved, we can do no better than estimate the effects of differential learning from what is known about these effects in other learning situations. The weak interference effect is not surprising from another point of view either, namely the relative dissimilarity among interfering items in stage II. When learning B–C association in stage II the backward B–A associations acquired in stage I must be unlearned. Presumably the A responses elicited by B stimuli are verbal representations of the non-verbal stimuli; and these verbal representations now interfere with the C responses which are numbers. If these C responses had been syllables or words it is very likely that the interference would have been stronger.

It has been shown in experiments where only verbal material has been used that the mediation effect is related to the meaningfulness of the mediator (Kjeldergaard, 1968). Evidence of such a relationship in mediation experiments where non-verbal material has been used seems to be lacking, but Price & Slive (1970) found that

'relevant' labels increased the recognition in a simple stimulus recognition paradigm where random shapes were used. In cases where there has been a failure to obtain the mediation effect (e.g. Barclay, 1961) this failure has been explained as partly due to the low degree of meaningfulness of the mediator. Thus it seems justified to assume that a higher degree of meaningfulness in the mediator should increase the mediation effect. It is not clear, however, how an increased meaningfulness would influence the asymmetry. It may be that there is an interaction between asymmetry and meaningfulness so that the asymmetric effects are more pronounced with high degrees of meaningfulness. Some earlier experiments reported by the authors (Björkman *et al.*, 1965; Garvill & Molander, 1968) give support to this assumption. In these experiments, the subjects had to a greater degree to rely on their own labelling of the stimuli and clear asymmetric effects were found. Fortunately, effects of meaningfulness and similarity among items can be investigated within the present approach.

Although the present experiment does not demonstrate that verbal mediation is necessary for cross-modal transfer to occur, it does demonstrate that verbal mediation is a powerful aid in transferring information between different sensory modalities, and that the effectiveness of the mediation may be different in different modalities. It also demonstrates that interference and mediation effects can be obtained in experiments where not entirely verbal material is used. Lastly, it suggests that the verbal mediation paradigms may be fruitful tools for a more systematic investigation of the role of verbalization in cross-modal transfer.

The authors are indebted to Dr B. Brehmer for valuable suggestions during the course of this work and Mr T. Krooks for collecting the data. The investigation was supported by a grant from the Swedish Council for Social Science Research.

References

Barclay, A. (1961). Objective mediators in paired-associate learning. *Am. J. Psychol.* **74**, 373–383.

Björkman, M. (1971). Cross-modal effects in human learning and judgment. *Proc. XIX Int. Congr. Psychol.* London: British Psychological Society.

Björkman, M., Garvill, J. & Molander, B. (1965). Cross-modal transfer as a function of preparatory set and distinctiveness of stimulus aspects. *Rep. from the Psychol. Lab. Univ. Stockh.* no. 186.

Blank, M., Altman, D. L. & Bridger, W. H. (1968). Cross-modal transfer of form discrimination in preschool children. *Psychon. Sci.* **10**, 51–52.

Blank, M. & Bridger, W. H. (1964). Cross-modal transfer in nursery-school children. *J. comp. physiol. Psychol.* **58**, 277–282.

Cashdan, S. (1968). Visual and haptic form discrimination under conditions of successive stimulation. *J. exp. Psychol.* **76**, 215–218.

Epstein, W. (1967). A Swedish evaluation of the association value of CVC nonsense syllables. (Unpublished manuscript.)

Ettlinger, G. (1967). Analysis of cross-modal effects and their relationship to language. In C. H. Millikan & F. L. Darley (eds.), *Brain Mechanisms underlying Speech and Language*. New York: Grune & Stratton.

Garvill, J. & Molander, B. (1968). Asymmetric effects in cross-modal transfer. *Umeå Psychol. Rep.* no. 3.

Gaydos, H. F. (1956). Intersensory transfer in the discrimination of form. *Am. J. Psychol.* **69**, 107–110.

Gibson, E. J. (1969). *Principles of Perceptual Learning and Development*. New York: Appleton-Century-Crofts.

Gibson, J. J. (1963). The useful dimensions of sensitivity. *Am. Psychol.* **18**, 1–15.

Horton, D. L. & Kjeldergaard, P. M. (1961). An experimental analysis of associative factors in mediated generalization. *Psychol. Monogr.* **75**, no. 11. (Whole no. 515.)

Houck, E. V., Gardner, D. B. & Ruhl, D. (1965). Effects of auditory and visual pretraining on performance in a tactile discrimination task. *Percept. mot. Skills* **20**, 1057–1063.

Jenkins, J. J. (1963). Mediated associations: paradigms and situations. In C. N. Cofer & B. S. Musgrave (eds.), *Verbal Behavior and Learning*. New York: McGraw Hill.

Jenkins, J. J. (1965). Comments on pseudomediation. *Psychon. Sci.* **2**, 97–98.

Jenkins, J. J. & Foss, D. J. (1965). An experimental analysis of pseudomediation. *Psychon. Sci.* **2**, 99–100.

Kjeldergaard, P. M. (1968). Transfer and mediation in verbal learning. In T. R. Dixon & D. L. Horton (eds.), *Verbal Behavior and General Behavior Theory*. Englewood Cliffs, N.J.: Prentice-Hall.

Lobb, H. (1965). Vision versus touch in form discrimination. *Can. J. Psychol.* **19**, 175–187.

Milne, A. M. (1968). A developmental study of touch and vision: form learning and cross-modal transfer. (Abstract of dissertation, George Washington University.)

Price, R. H. & Slive, A. B. (1970). Verbal processes in shape recognition. *J. exp. Psychol.* **83**, 373–379.

Rock, I. & Harris, C. S. (1967). Vision and touch. *Sci. Am.* **216**, 96–104.

Sinha, A. K. P. & Sinha, S. N. (1960). Intersensory transfer in learning sequences. *J. exp. Psychol.* **60**, 180–182.

Underwood, B. J. (1964). Degree of learning and the measurement of forgetting. *J. verb. Learn. verb. Behav.* **3**, 112–129.

Walk, R. D. (1965). Tactual and visual learning of forms differing in degree of symmetry. *Psychon. Sci.* **2**, 93–94.

Wright, J. M. von (1970). Cross-modal transfer and sensory equivalence – a review. *Scand. J. Psychol.* **11**, 21–30.

(*Manuscript received* 4 *December* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 459–468
Printed in Great Britain

WHY DO CHILDREN REVERSE LETTERS?

By UTA FRITH

MRC Developmental Psychology Unit, Gordon Street, London

When normal letters and their mirror-image forms were presented for copying and matching, school children made more reversal errors with mirror-image forms than with normal letters. Preschool children made equally many reversal errors with both types of material. This was also true for older children when unfamiliar letter-like symbols were presented. The results are described in terms of discriminability and response bias as derived from signal detection analysis. Keeping discriminability constant, a significant increase in response bias was shown by 5 year-olds as compared to 4-year-olds, due presumably to the effect of one year's normal school practice. It is hypothesized that children showing persistent reversal errors fail to acquire a preference for the given spatial orientation of letters as rapidly as normal readers.

The observation that children learning to read and write tend to reverse letters and numbers, e.g. И for N, has become a preoccupation with many investigators of perceptual learning. Since errors of this kind appear to persist after the age of eight in children with reading and writing difficulties, and since they are also found in people suffering from certain brain injuries affecting parietal and occipital regions (Benton, 1962; Luria, 1965), such errors have been considered an important diagnostic sign. In theories of dyslexia in particular, right–left reversals have been considered characteristic of the disorder (Orton, 1937; Hermann, 1959; Zangwill, 1960).

Mirror reversals and rotations in the same plane have not always been clearly separated, and most theories try to account for both types of errors at the same time. In fact, both errors can be considered as relating to an uncertainty about the spatial orientation of shapes. Neurological, maturational and psychological theories have been employed to explain the phenomenon and an extensive discussion has been provided by Fellows (1968). He stresses that several factors together have to be taken into account, such as discrimination ability, attention, learning and memory. A theory placing much emphasis on the perceptual aspects in overcoming reading and writing errors is, for example, the theory proposed by Gibson *et al.* (1962). The ability to discriminate between two shapes is doubtless an essential prerequisite in order to learn to read. Gibson *et al.* (1962) and Hendrickson & Muehl (1962) demonstrated that children have to learn to attend to the feature of orientation which is critical for differentiating letters. The left–right and up–down orientation (amongst other distinctive features) of a graphic symbol has to be recognized by the child as important. This is by no means obvious, since the dimension of orientation is irrelevant in identifying three-dimensional objects. While it may be of little consequence if very young children at first ignore orientation, it is necessary for them to be aware of this feature in order to handle two-dimensional shapes.

A theory emphasizing the role of learning or memory in learning to read and write is expressed, for example, by Schubenz & Buchwald (1964). They found that errors in writing single letters were related to the frequency of occurrence of the letter in the German language. Thus they concluded that children who show frequent reversal and other errors suffer from a memory deficit. Frequently occurring letters are

practised more and remembered better and are thus less likely to be produced incorrectly.

Using a quite different method, Krise (1952) also found that learning factors were of utmost important in relation to reversal errors. He demonstrated that normal adults show a strong reversal tendency if presented with unfamiliar symbols analogous to 'b' and 'd'. He concluded that the tendency to reversal is present in everybody, perhaps due to some physiological weakness, and can only be overcome or inhibited by specific learning.

The following experiments were designed to assess the role of both perceptual and learning factors in order to explain why children reverse letters. Reversal errors might be due to an inability to perceive the difference between a shape and its mirror-reversed form. This inability might be similar to that of a tone-deaf person, when required to discriminate between different tones. On the other hand, errors might be due to incorrect memory of the letter's orientation in space. This might be independent of whether or not a shape could be discriminated from its mirror image.

In order to test these hypotheses, a method was used which required the presentation of letters in their normal form and in their mirror-reversed form. Each of the forms had to be copied or compared to a standard. An error of reversal in the first case would result in an unusual orientation, while in the second case such an error would result in the usual and familiar orientation of the letter. If errors would occur equally often in either case, then the assumption of difficulties in shape discrimination would be sufficient. If errors were more frequent in relation to mirror-reversed shapes (i.e. when the error resulted in the usual orientation of a letter), then something like a bias for the left–right orientation of shapes would have to be postulated.

An underlying assumption for these experiments should be made explicit: it is assumed that spatial orientation can be separated from shape and that one can talk about the basic shape in a pair of stimuli, e.g. a letter and its mirror image, regardless of left–right orientation. The hypothesis of a bias for orientation would imply that this dimension is stored and coded separately from other shape characteristics and that the code may be strong enough to override 'perception'.

EXPERIMENT I

Copying shapes

Method

In a study carried out by the author (Aurnhammer, 1966) 215 London school children, aged 4–9 years, were tested on a test battery which included copying letters and numbers which were presented in normal as well as mirror-reversed form. The 20 letters and digit pairs that were used are shown in Table 1. The symbols were produced by means of Uno penstencil, UL, pen 0·7, the size of the symbols being $\frac{3}{16}$ in. ($+\frac{1}{10}$ in. for lower for upper lines in letters like 'y'). These symbols were presented in random order, first in their normal, and afterwards, on a separate sheet, in their mirror-reversed form. The objection might be raised that this introduced a systematic order effect. However, since the children at school had the usual practice in writing normal letters, a set for the normal spatial orientation should have been already established and, in relation to this effect, the order effect should have been negligible.

The children had to copy each symbol beside it on the same sheet of paper. This was done as a group test. The instruction was to copy the symbols as exactly as possible. All errors of reversal and rotation were counted. It was observed that mirror reversals were more common for these symbols than rotations, so that, although the present score confounds both error types, rotations played only a negligible part.

The letters b, d, p, q, all of which are reported to be particularly prone to reversal errors, were not included in this task, since the distinction between mirror-reversed form and normal form becomes meaningless. If a mirror image would be presented for copying, it would be at the same time the normal form of another letter.

Results

Errors of copying were obtained for the normal form and the mirror-reversed form of each shape. An error score indicates the number of children out of 215 who reversed or rotated either form of a basic shape. The results are shown in Table 1. It can be seen that relatively few errors occurred in most cases and that errors tended to be more frequent when the mirror-reversed form of a stimulus pair was to be copied. In order to evaluate to what extent this was due to a response bias for the familiar left–right orientation of the shapes, and in order to compare it with the discriminability of the shapes, the following analysis, developed by Luce (1963), was carried out. Two scores, η and b in Luce's terminology, were calculated. To do this, the frequency of correct and incorrect copies for each pair of shapes was set out in tables as in the example below:

		Stimulus presented	
		Mirror И	Normal N
Response obtained:	Incorrect	N = 43 (A)	И = 10 (B)
	Correct	И = 172 (C)	N = 205 (D)

To calculate the discriminability of a stimulus pair (e.g. N–И) the ratio of incorrect to correct copies was obtained. This gives an estimate of the error-proneness of a stimulus pair, regardless of which of the two forms had to be copied. It is preferable to the total number of errors as an index, since it is independent of the frequency with which it was presented or responded to. Thus a score $\eta = \sqrt{(AB/CD)}$ is available for each pair. A score of 0 designates maximum discriminability, while a score of 1 would indicate a complete lack of discriminability.

In addition, for all stimulus pairs a score for the strength of the response bias was obtained, in order to see how often one particular left–right orientation had been reproduced in preference to the other, regardless of correctness. This score is calculated as follows: $b = \sqrt{(AD/BC)}$. A score greater than 1 would indicate that the response tended to be more often of the normal form of the letter, while a score smaller than 1 would indicate that the response tended to be more often the mirror-reversed form.

Table 1 contains the η and b scores for all letter and digit pairs. They are presented in such a way that the stimulus pairs are categorized according to the size of their scores. Arbitrarily a dichotomy was applied so that the 10 pairs with the lowest b scores were categorized under weaker bias and the 10 pairs with the highest b scores were categorized under stronger bias. Similarly, the 20 stimulus pairs were divided into two groups of lower and higher discriminability although all η scores were near 0 and thus indicated relatively high discriminability.

It must be pointed out that the scores are less valid for the stimulus pairs based

Table 1. *Copying errors, discriminability (η) and response bias (b) for normal letters and numbers and their mirror image*

(The 20 stimulus pairs are classified in terms of the size of η and b scores, so that the 10 highest scores appear in a separate category from the 10 lowest scores. Errors indicate the number of children out of 215 who reversed the stimulus when copying it.)

Lower discriminability								Higher discriminability							
Stronger bias				Weaker bias				Stronger bias				Weaker bias			
	Error	η	b		Error	η	b		Error	η	b		Error	η	b
n n	5 58	0·09	3·94	Z Ꙁ	11 12	0·06	1·05	y γ	0 11	0·01	3·40	G Ә	1 2	0·00	1·42
N И	10 43	0·11	2·26	4 ᔭ	8 9	0·04	1·06	h ʮ	0 8	0·01	2·88	K ꓘ	0 2	0·00	1·42
S Ƨ	1 21	0·02	4·80	7 Ⲅ	3 14	0·03	2·22	B ꓭ	0 7	0·01	2·68	C Ɔ	0 2	0·00	1·42
e ɘ	0 22	0·02	4·93	a ɒ	3 6	0·02	1·42	6 ∂	0 7	0·01	2·48	E Ǝ	0 1	0·00	1·00
5 ꙅ	1 16	0·02	4·15	9 ɘ	5 4	0·02	0·89	u u	0 6	0·01	2·48	F ꟻ	0 1	0·00	1·00

on E, F, C, K, G, since they were associated with extremely few, i.e. one to three errors. In cases where no error occurred in one condition, a score of 1 had to be substituted in the zero cell of the contingency table, since otherwise the calculation would have included a division by 0. For this reason the scores for y, h, B, 6, u, might also be of restricted validity.

Discussion

The errors indicated that there was a tendency to reproduce the normal form of a letter or number whatever orientation the shape was presented in. Thus reversal-proneness of a shape may be quite different according to which spatial orientation it is presented in. Signal-detection analysis enabled a quantitative description of two factors involved in the task: first, response bias, which was indicated by the greater amount of errors in relation to a mirror-reversed symbol, and secondly, discriminability, which is related to some shapes being more difficult than others. Thus the letters and numbers could be categorized according to what extent they were associated with a bias regardless of the total number of errors and to what extent they differed in terms of discriminability. In this particular group of children, the stimulus pairs were on the whole of very high discriminability and differed little on this measure, while some stimuli showed a much stronger response bias than others.

Taking only the number of errors would have been quite misleading in elucidating these factors. For example, although the pair 'n–ᴨ' obtained many errors, this does not imply that it is of particularly low discriminability. Few errors occurred when the basic shape was presented in its normally encountered form. This was presumably due to the strong response bias associated with this stimulus pair, so that the normal form was produced, whichever of the two orientations had been presented for copying. It may perhaps be suggested that the relatively large number of errors occurred as a result of oversight, since in this particular stimulus only a minor detail serves as an anchor for the spatial orientation. However, the 'oversight' was almost 12 times as frequent when the mirror image was presented. In comparison, the very similar looking letter pair 'u–ᴝ' did not suffer from such 'oversight'. The bias associated with this pair was not as strong as for 'n–ᴨ'. This may be because 'u' occurs less frequently in normal usage than 'n'.

Rough estimates of the frequency of letters in English suggest that letters associated with a stronger bias tended to be more frequent than letters with a weaker bias. There was a difference between an average frequency of 5·8 and 3·9 per cent. This finding would be consistent with the hypothesis that the learning of letters is normally associated with the build-up of a response bias, so that one orientation rather than another of a particular shape becomes preferred. In the present group children of a wide age range and thus of different levels of reading achievement were included. Nevertheless, it could be demonstrated that a bias exists towards the normal and familiar orientation of letters and digits. It seems probable that the inclusion of younger children with little reading and writing practice has led to an underestimation of this effect, since these children could not be expected to have learned the familiar orientation of letters and numbers as well as older children. Unfortunately, it was not feasible to subdivide the results according to age-groups, since, due to the low error rate, this would have resulted in unreliability.

The frequency of usage of letters was also related to the degree of discriminability. Somewhat surprisingly, it appeared that letters with lower discriminability tended to be more frequent than letters with higher discriminability (6·2 v. 3·4 per cent average frequency of occurrence). Taking into account the very small differences in terms of discriminability found, one may conclude that the frequency of a letter in the English language influences the build-up of a response bias, but does not relate to discriminability. By inference, if practice is largely determined by frequency of occurrence, learning to code the spatial orientation of letters may have little to do with an improvement in perceptual discrimination.

There was one exception to the general tendency for a bias towards the usual orientation of a symbol, namely the digit pair '9–e', where there was a slight preference for the mirror image. The mirror image of '9' bears considerable resemblance to the letter 'e' in hand-writing. Since 'e' is the most frequently used letter of the alphabet, it can be assumed that its particular left–right orientation is excessively well practised, so that a strong bias might well exist in its favour.

The particularly well-known and persistent reversal errors for the letters b–d and p–q may be caused not so much by a lack of discriminability but by the lack of a bias for one specific orientation of the basic shape. Different orientations of the basic shape in question are probably encountered equally frequently and thus a response bias is less easily established. It is plausible that the coding in such a case is more difficult since it must provide for the conditions 'if *q*, move to left' and 'if *p*, move to right' in relation to the same basic shape, while another basic shape needing the code 'move to left' only would perhaps more easily lead to a response bias.

The purpose of the first experiment was to see whether a response bias could be demonstrated in young school children. There was sufficient evidence for this to warrant the second experiment which investigated the effect of age or, more precisely, the effect of reading and writing practice as gained through normal school experience. In this experiment only such shapes were used where a particularly large error rate had been established in pilot studies so that the reliability of a signal detection analysis would be increased.

EXPERIMENT II

Matching shapes

Two groups of children were compared: one younger group, unfamiliar with letters, and one older group after one year at school. It was found out in a pretest that the younger children were not able to copy symbols. Thus the task was to match symbols, which were again presented in two forms, as 'normal' and as mirror-reversed form of the same basic shape.

If response bias depended on practice gained at school, then the younger children should show no response bias in relation to the letter pairs. This should also be the case for the older children, when unfamiliar shapes are presented.

Method

Subjects. Ten children of an average age of 4:3 years (range 3:5 to 4:6) who visited a nursery school, and who could neither read nor write letters, made up the younger age-group. The older age-group was made up of 10 school children of an average age of 5:10 (range 5:4 to 6:2) years. Their intelligence was not tested, but according to teachers' estimates the children were either

average or slightly below average in attainment. The two groups were matched individually in terms of total errors made on the experimental task. On average both groups had an attainment level of approximately 60 per cent correct responses. The data on age and overall performance are included in Table 2.

In order to achieve individual matching across groups, 10 more children were tested than were actually used in the analysis of results. They were excluded when only the total number of errors was known, so that no systematic selection could have taken place. In order to avoid ceiling or floor effects all children outside the error range from 33 to 66 per cent were rejected. By controlling the level of performance it is ascertained that any interaction effect between groups and conditions cannot be attributed to a difference in task accomplishment. In the present experiment matching in terms of total errors results in controlling the factor 'discriminability' so that the factor 'bias', which constitutes the main object of the investigation, can be studied independently from contamination by differential error rate.

The question arises whether older children who made the same number of errors on the given task were less intelligent than the younger children, but fortunately this factor is not relevant for the purpose of the present experiment.

Material. The material consisted of three letters (N, J, Z), three numbers (5, 7, 9) and six similarly constructed nonsense symbols (Ꮛ, √, ∡, ʃ, ⊌, ♆). Among the latter were some of Gibson's (1962) letter-like shapes. The particular letters and numbers were chosen because in a pilot study they were found to be particularly prone to errors of rotation and reversal. The letter-like shapes were selected so as to be relatively similar to the letters and numbers used.

There were 12 stimulus pairs, one of each being the mirror image of the other. Three cards (6 × 8 in.) were prepared with four of these stimulus pairs on each card. There were also standard shapes, on single cards (2 sq. in.), which were identical with the 24 stimuli ($\frac{4}{16}$ in. high, $\frac{3}{16}$ in. wide).

Procedure. There was a short pretraining session in which an example was explained and demonstrated. The stimulus pair 'S' 'Ƨ' was used for this. The instructions were: 'These two look alike, but they are not the same. This one goes this way and that one goes that way: it goes the *other* way round.' The child was then made to trace the direction of the two stimuli with his finger. Then a standard ('S') was presented and the child had to point to the form which showed exactly the same direction. Only when the child had passed two trials with two standards ('S' and 'Ƨ') was the experiment started.

A standard shape was presented and its match had to be chosen from eight symbols on the card. The standard was immediately in front of the child and in front of the eight symbols. The child had to point to the correct stimulus. 'Show me, which one is exactly the same? Which one goes the same way?' was the instruction used. Errors of mistaking one basic shape for another, e.g. a 5 for a J, were avoided insofar as on rare occasions where they occurred, the child was given another trial. Errors of mismatching in terms of the directions of a stimulus were recorded.

Results

On average the children mismatched 10 out of 24 stimuli. All children made on the whole significantly more errors with letter-like shapes than with familiar letters and numbers ($t = 2{\cdot}60$; $P = < 0{\cdot}02$). The results are shown in Table 2. There was no interaction involving a group difference, which might indicate that even the younger children could somehow distinguish real letters and numbers better than nonsense symbols. This suggests that they were perhaps not entirely unfamiliar with written material. The matching of the groups precluded any difference in the total number of errors. However, group differences became apparent in the analysis of errors relating to letters and numbers only. Means and S.D.s of these errors are shown in Table 3. The older children made significantly more errors when the standard was a mirror image of a letter or number ($t = 2{\cdot}59$; $P < 0{\cdot}02$). On the other hand, for the younger children there was no significant difference between errors in the two conditions.

The errors over all familiar stimulus pairs could be used to calculate discriminability

(η) and bias (b) analogous to the analysis of the first study. However, whereas the scores in the previous study were obtained for each individual stimulus pair over all subjects, the scores in the present study were obtained for each individual subject over all letters and numbers in order to allow group comparisons.

Table 2. *Errors of matching familiar and unfamiliar shapes*

Groups	Age (months)	Total errors (max. = 24)	Errors with familiar letters and numbers (max. = 12)	Errors with unfamiliar letter-like shapes (max. = 12)
Younger children				
Mean	49·90	10·30	4·70	5·60
S.D.	3·84	2·54	1·57	1·43
Older children				
Mean	69·70	10·30	4·30	6·00
S.D.	2·87	2·31	1·89	1·83

Table 3. *Errors of matching familiar letters and numbers* (max. = 6)

Groups	Stimulus presented: Normal orientation	Mirror image
Younger children		
Mean	2·20 (37 %)	2·50 (41 %)
S.D.	0·79	1·08
Older children		
Mean	1·70 (28 %)	2·60 (43 %)
S.D.	0·82	1·35

Table 4. *Discriminability and response bias based on familiar letters and numbers*

Groups	Discriminability η (log) (score of 0 = chance performance)	Bias b (log) (score of 0 = bias absent)
Younger children		
Mean	−0·27	0·00
S.D.	0·33	0·18
Older children		
Mean	−0·25	0·14
S.D.	0·26	0·15

The η and b scores were transformed to logarithms to normalize the distribution and also to make them equivalent to scores of detectability as used in signal detection theory. Table 4 shows means and S.D.s. On these logarithmic scores t tests were carried out to compare younger and older children. In both groups η scores (discriminability) were significantly different from chance ($t = 2·55$; $t = 2·65$; $P < 0·05$). There was no group difference in terms of η. However, the b scores (bias) indicated a significant bias in the older group only ($t = 2·99$; $P < 0·02$), and the two groups were significantly different from each other on this measure ($t = 2·54$; $P = 0·02$).

In error scores for unfamiliar, letter-like symbols, there was no indication of bias or change of bias with age.

Discussion and conclusions

The second experiment showed that even if matching rather than writing is required, the use of discriminability and response bias [illegible] children to reveal errors. Thus, response bias need not be thought of in terms of firmly engrained hand movements as a result of writing practice.

Keeping discriminability constant, which was effectively achieved by matching the two groups of children in terms of total performance, a response bias was clearly demonstrable in the older group but not in the younger one. This means that older children made errors when familiar letters and numbers were concerned mainly because of a strong preference for the familiar orientation of the shape. There was no such bias concerning unfamiliar shapes and there was no bias in the preschool children, who were unfamiliar with all the stimuli. Thus one can conclude that as little as one year of exposure to letters was sufficient to build up a bias for orientation in normal children. This conclusion holds even if one wanted to assume that the older children were less intelligent than the younger ones, on the basis that they made the same total number of errors. However, the relationship between response bias acquisition and intelligence may be an interesting problem for investigation.

The results obtained in the older children are an illustration of Garner's (1966) claim "To perceive is to know". Even though simultaneous comparison was available the act of matching was sometimes more strongly governed by the knowledge of the normal orientation of the letter. Thus the correct match was chosen when the external stimulus coincided with the internal representation and the incorrect match when the external stimulus did not coincide with it. In fact, one might say that some children behaved as if matching to the internal standard and not to the external one. The present results surely are an example of the strength of internally represented knowledge (here described in terms of a response bias) over the presumed perception of external stimuli.

It could be hypothesized that a certain subgroup of children with exceptional reading and writing difficulties might be incapable of forming a bias for left-right orientation. If this were true, their handicap could not be defined in terms of a perceptual dysfunction, but rather in terms of a specific learning impairment.

Many investigators have stressed that severe reading retardation is often associated with difficulties not confined to the writing of symbols, particularly a lability of discriminating left and right (Benton, 1962; Vernon, 1962; Rutter *et al.*, 1966; Lyle, 1969). This difficulty could be understood in terms of a lack of bias as defined in the experiments. Relevant to this hypothesis is the report of Kohler's (1951) about the effects of wearing right–left inverting prisms. Smith & Smith (1962) summarized the results, of which the following seem to be of particular relevance. After wearing the prisms for only 3–4 weeks, a subject did not see letters as reversed any more. Removal of the prisms after some initial disorientation again quickly reinstated the previous bias. On the other hand, writing was never affected by the reversal of vision, presumably since preferences for particular orientations of letters had become automatic in writing movements. Thus it seems that orientational bias, once established in movement, i.e. largely independent of vision, cannot easily be disrupted by

visual rearrangement of spatial relations. On the other hand, a new bias in the visual perception of spatial orientation can apparently be very rapidly acquired.

It could be hypothesized that people with the type of severe reading and writing difficulties, which are due to a lack of memory for spatial orientation, would not readjust quickly to inverting spectacles. The tendency to getting rapidly and strongly adjusted to right and left orientations might distinguish normal from backward readers.

While a difficulty in discriminating right and left plays a significant role in some forms of dyslexia, it is nevertheless recognized that this difficulty decreases with increasing age. It is also known that backward readers do not usually show gross disorientation in space or apraxic difficulties. The theory that a failure to acquire orientation preferences is responsible for reversal errors is compatible with the notion that prolonged learning will eventually establish such a preference. While normal children acquire the bias very quickly, some dyslexic children might acquire it only very slowly. Thus activities which are highly practised would not show up this deficit, and normal and backward readers might only be distinguished at early stages of learning.

References

AURNHAMMER, U. (1966). Rotations and reversals in children's writing. (Unpublished dissertation, University of London.)

BENTON, A. L. (1962). Dyslexia in relation to form perception and directional sense. In J. Money (ed.), *Reading Disability*. Baltimore: Johns Hopkins Press.

FELLOWS, B. (1968). *The Discrimination Process and Development*. Oxford: Pergamon Press.

GARNER, W. R. (1966). To perceive is to know. *Am. Psychol.* **21**, 11–19.

GIBSON, E. J., GIBSON, J. J., PICK, A. D. & OSSER, H. (1962). A developmental study of the discrimination of letter-like forms. *J. comp. physiol. Psychol.* **55**, 897–906.

HENDRICKSON, L. N. & MUEHL, S. (1962). The effect of attention and motor response pretraining on learning to discriminate *b* and *d* in kindergarten children. *J. educ. Psychol.* **53**, 236–241.

HERMANN, K. (1959). *Reading Disability*. Copenhagen: Munksgaard.

KOHLER, I. (1951). Über Aufbau und Wandlungen der Wahrnehmungswelt insbesondere über bedingte Empfindungen. *Oest. Akad. Wiss. phil. hist. Klasse Sitzungsber.* **227**, 1–118.

KRISE, M. (1952). An experimental investigation of theories of reversals in reading. *J. educ. Psychol.* **43**, 408–422.

LUCE, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush & E. Galanter (eds.), *Handbook of Mathematical Psychology*, vol. 1. New York: Wiley.

LURIA, A. R. (1965). Neuropsychological analysis of focal brain lesions. In B. B. Wolman (ed.), *Handbook of Clinical Psychology*. New York: McGraw-Hill.

LYLE, J. G. (1969). Reading retardation and reversal tendency: a factorial study. *Child Dev.* **40**, 833–843.

ORTON, S. T. (1937). *Reading, Writing and Speech Problems in Children*. New York: Norton.

RUTTER, M., YULE, W., TIZARD, J. & GRAHAM, P. (1966). Severe reading retardation: its relationship to maladjustment, epilepsy and neurological disorders. *Proc. First Int. Conf. Assoc. for Special Education*. Stanmore: Assoc. for Special Education.

SCHUBENZ, S. & BUCHWALD, R. (1964). Untersuchungen zur Legasthenie. I. Die Beziehung der Legasthenie zur Auftretenshäufigkeit der Buchstaben des Alphabets in der Deutschen Sprache. *Z. exp. angew. Psychol.* **11**, 155–168.

SMITH, K. U. & SMITH, W. M. (1962). *Perception and Motion*. Philadelphia: Saunders.

VERNON, M. D. (1962). Specific dyslexia. *Br. J. educ. Psychol.* **32**, 143–150.

ZANGWILL, O. L. (1960). *Cerebral Dominance and its Relation to Psychological Function*. Edinburgh: Oliver & Boyd.

(*Manuscript received* 11 *July* 1970; *revised manuscript received* 7 *December* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 469–473
Printed in Great Britain

EXAMINATION AND SEARCH IN INFANTS

By P. L. HARRIS

Department of Experimental Psychology, University of Oxford

[illegible] of 'search' v. 'no search'.

Piaget (1954) has argued that the ability to search for a hidden object is of central importance in early cognitive development. He suggests that search by the infant indicates that the object is credited with a continuity or permanence despite its disappearance from the visual field, in that it is still regarded as available to the hand.

In a longitudinal study, Schofield & Uzgiris (1969) investigated possible preconditions for the appearance of search. They found that examination of an object (i.e. coordinated visual and manual inspection) increased in the weeks prior to the development of search. There is, however, no experimental evidence for a relationship between examination and search.

Examination of an object might facilitate search in that it provides manual contact with the object. Piaget (1954) argues that at first the infant will only search for an object he has touched immediately prior to its disappearance. Manual search following visual inspection alone appears later. In addition to providing manual contact, examination of an object might facilitate search by permitting a longer duration of exposure to the object prior to its disappearance.

The following experiment was designed to investigate the role of the two variables discussed above, manual contact and duration of examination. Age differences were also investigated by comparing infants from 7 to 9½ months with infants from 11 to 13½ months. Half the subjects in each age-group were presented with the object in a Perspex box; this allowed visual inspection only. Half were allowed both visual and manual inspection of the object. Each subject was given an opportunity to search following each of three different periods of examination.

METHOD

Subjects. The subjects were 24 infants aged 11–13½ months and 24 infants aged 7–9½ months. Each age-group was equally subdivided into a visual subgroup (V) and a visual + manual subgroup (V + M). In the younger group the mean age was 247 days (S.D. 17) for the V subgroup and 262 days (S.D. 26) for the V + M subgroup. In the older group the mean age was 358 days (S.D. 18) for the V subgroup and 369 days (S.D. 18) for the V + M subgroup.

Apparatus and stimuli. A screen, painted matt-black, was placed on the table between the subject and the experimenter. A white box was placed flush against the screen on the subject's side.

An aperture in the screen provided the experimenter with access to the [illegible] at the back. A wooden door, faced with leather material, was hinged to the [illegible] the box. This allowed passage to objects emerging from the box towards the [illegible] tion. The infant could push the door inwards and upwards, when attempting to [illegible] object which disappeared into the box. When the door was displaced [illegible] switch was activated which was wired to a Rustrak event recorder. This [illegible] when the door was displaced until the infant withdrew his hand allowing [illegible] its original position. This made it possible to measure the persistence of [illegible]

Objects were conveyed to the infant for examination by means of a [illegible] Fig. 1). This ran from a position directly in front of the subject back to the [illegible] The experimenter could therefore place an object inside the box and on the [illegible]

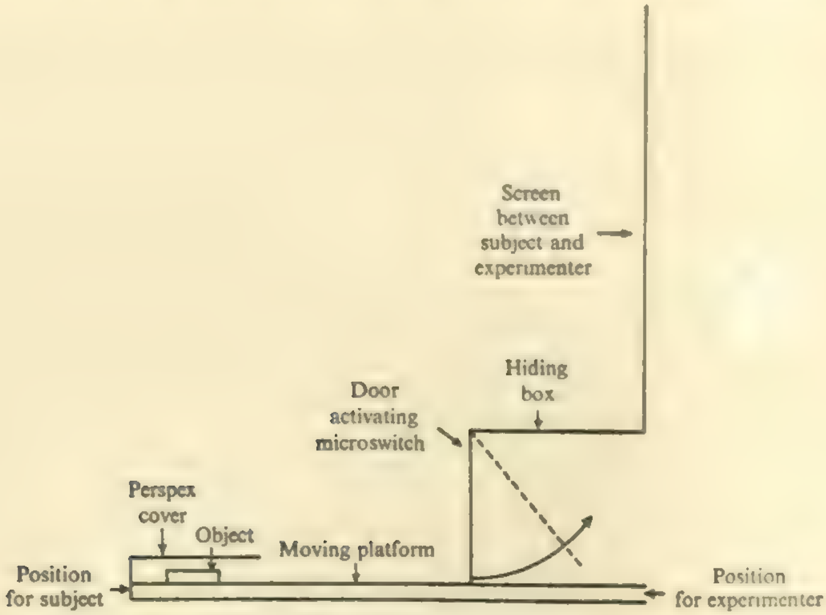


Fig. 1. Schematic diagram of apparatus (not drawn to scale).

position behind the screen. The object could then be pushed along the platform by the experimenter until it emerged from the box and reached the end of the platform nearest the infant. For subjects in the V subgroup the object was moved along the platform into a Perspex cover which prevented manual access to the object. For subjects in the V + M subgroup the front portion of the Perspex cover was removed, allowing manual access. The back portion was retained so that, when the object disappeared back into the box, the infant was forced to release the object as it moved underneath the Perspex.

Three different objects were used: a stopwatch, a bunch of keys and an (empty) red razor-blade dispenser. All objects were fixed to the moving platform.

Procedure. Apart from age, the major between-subjects variable was mode of examination, i.e. visual versus visual + manual.

The major within-subjects variable was duration of presentation. Each subject received three trials: one of 5 sec., one of 20 sec. and one of 80 sec. duration.

All subjects received a different object for each of the trials. The order of presentation of the objects and the order of the three trials with respect to the duration of presentation were counterbalanced.

Each subject was seated at the table on his mother's lap. The procedure of the experiment was explained to the mother. She was asked to do nothing except hold the infant so that he could retrieve the object by pushing the door. (This procedure, rather than the customary use of

[illegible]

Table 1. *Mean persistence of search in seconds (transformed scores) as a function of age, condition (visual versus visual + manual) and duration of presentation*

	5 sec.	20 sec.	30 sec.
Older			
Visual	2·10	2·91	5·70
Visual + manual	2·35	4·35	3·14
Younger			
Visual	0·15	0·10	0·26
Visual + manual	1·70	3·28	1·77

Table 2

(a) Final table of the total analysis of ages (A) × duration (B) × mode of examination (C)

Source	M.S.	D.F.	F	P
A	4·329	1	19·034	0·001
C	1·574	1	6·920	0·025
AC	1·807	1	7·945	0·01
Subjects within groups	0·227	44	—	—
B	0·225	2	2·005	—
AB	0·105	2	0·938	—
BC	0·266	2	2·375	—
ABC	0·032	2	0·573	—
B × subjects within groups	0·112	88	—	—

(b) Further analysis of ages (A) × mode of examination (C) interaction

Source	M.S.	D.F.	F	P
A at C_1	5·865	1	25·838	0·001
A at C_2	0·270	1	1·190	—
C at A_1	0·004	1	0·017	—
C at A_2	3·377	1	14·878	0·001
Subjects within groups	0·227	44	—	—

A_1, older; A_2, younger; C_1, visual; C_2, visual + manual.

Results

In order to reduce heterogeneity of variance the scores were subjected to a logarithmic transformation ($x = \log_{10} x + 1$) (Kirk, 1968). Table 1 presents the transformed mean scores for the persistence of search as a function of age, mode of examination and duration of examination.

The results were subjected to analysis of variance. Preliminary analysis established

that neither objects nor order of duration were significant, either as main effects or in interaction with other variables.

The final table for the analysis of variance is presented in Table 2. The significant results were a difference in ages ($P < 0{\cdot}001$) and mode of examination ($P < 0{\cdot}025$) and the interaction of age × mode of examination ($P < 0{\cdot}01$). Duration had no significant effect either as a main term or in interaction with the other two variables.

The ages term showed that the older group searched more persistently than the younger group. The mode of examination term showed that search was more persistent in the visual + manual subgroups than in the visual subgroups.

However, both effects must be qualified by reference to the interaction of age × mode of examination; further analysis of this interaction (Kirk, 1968) is presented in Table 2*b*.

The older group searched more persistently than the younger group in the visual condition (A at C_1, $P < 0{\cdot}001$) but there was no significant difference between ages in the visual + manual condition (A at C_2, P n.s.). This further analysis also showed that the younger group searched more persistently in the visual + manual condition than in the visual condition (C at A_2, $P < 0{\cdot}001$). The older group, however, searched no more persistently in the visual + manual condition (C at A_1, P n.s.).

In summary, if the object was handled the younger infants searched as persistently as older infants. If the object was only seen prior to its disappearance the younger group searched much less persistently, if at all. Persistence of search did not appear to be affected by duration of examination.

Discussion

In the introduction two questions were asked about the effect of examination of an object on search for that object. Firstly, it was asked whether examination facilitates search in that it provides manual contact. Secondly, it was asked whether longer periods of examination might facilitate registration and recall of the object, thereby leading to more persistent search.

The results of the experiment give a partial reply to both of these questions. Manual contact appears to be important in that the visual + manual condition led to more persistent search than the visual condition, but only for the younger group. The duration of examination of the object did not lead to more persistent search.

Manual contact appears to serve as a prime stimulus for the elicitation of search in the younger group. In the older group visual inspection alone can serve the same function. It may be asked exactly how manual contact serves as a prime stimulus, and how it comes to be redundant for older subjects.

Piaget emphasizes the role of manual contact, since he argues that search is initially merely an extension of reaching movements. An infant who has touched and therefore reached for an object is more likely to extend his hand in search when it disappears than an infant who has only engaged in visual inspection. This hypothesis seems unlikely to account for the impairment of the younger subjects in the visual subgroup of this experiment. This is because, although subjects in this subgroup could not touch the object, they could still make reaching movements. Informal observation suggested that the majority of subjects in the visual subgroup engaged

in such movements. Further experimentation could establish whether their number and duration differed in the visual subgroup as compared with the visual + manual subgroup. A second possibility arises from the presence of the Perspex box in the visual condition. Since this differed in shape from the object and remained behind when the object disappeared, it may have distracted younger infants from the object itself. Older infants, on the other hand, may have been able to ignore this distraction. However, in the visual + manual condition the back-portion of the Perspex box was retained, so that as the object passed under it, the infant was forced to release the object. Since this also differed in shape from the object and remained behind when the object disappeared, it would also constitute a distraction. Yet the younger infants searched as persistently as the older infants in the visual + manual condition.

A further alternative to the two suggestions considered above is to make a distinction between the visual and manual systems and to suggest that since search involves the manual system, it may require a manual prime stimulus. In younger infants this can only be provided when manual contact accompanies visual inspection; in older infants, who may have developed some form of cross-modal ability in the meantime, the visual stimulus can be translated into a manual stimulus.

It should be noted that the development of such an ability has also been postulated by Schaffer & Parry (1969). They argue that whilst infants as young as 6 months show visual discrimination of a novel from a familiar object, this information is not used to guide manual approach until the final quarter of the first year of life.

The experiment demonstrates that persistence of search in infants can be meaningfully related to other experimental variables. This measure has clear advantages over a simple binary classification of search versus no search, both in terms of the objective recording of responses and in terms of their statistical treatment thereafter.

This work was supported by an M.R.C. studentship. The author would like to thank Dr P. E. Bryant for his supervision and Dr K. Hope for his statistical advice. The cooperation of the Oxford Health Department and the health visitors of child welfare clinics at Blackbird Leys, Donnington Bridge, Banbury Road and East Oxford is gratefully acknowledged.

References

Kirk, R. E. (1968). *Experimental Design: Procedures for the Behavioural Sciences.* Belmont, Calif.: Brooks/Cole.

Piaget, J. (1954). *The Construction of Reality in the Child.* New York: Basic Books.

Schaffer, H. R. & Parry, M. H. (1969). Perceptual-motor behaviour in infancy as a function of age and stimulus familiarity. *Br. J. Psychol.* **60**, 1–9.

Schofield, L. & Uzgiris, I. C. (1969). Examining behaviour and the development of the concept of object. (Paper presented to the Society for Research in Child Development, Santa Monica.)

(*Manuscript received* 19 *August* 1970; *revised manuscript received* 2 *October* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 475–479
Printed in Great Britain

SOME FINDINGS RELEVANT TO THE HYPOTHESIS THAT TOPOLOGICAL SPATIAL FEATURES ARE DIFFERENTIATED PRIOR TO EUCLIDEAN FEATURES DURING GROWTH

BY DONALD COUSINS AND EUGENE ABRAVANEL
Department of Psychology, The George Washington University, Washington D.C., U.S.A.

The Piaget & Inhelder hypothesis that topological spatial features are differentiated prior to Euclidean features ontogenetically was tested with samples of children between $3\frac{1}{2}$ and $5\frac{1}{2}$ years. On a comparative similarity task, subjects indicated which of two comparison figures differing in topological and Euclidean features looked 'most like' a standard figure possessing some features in common with either comparison. At all ages, the majority of similarity judgements were in terms of Euclidean features. This was especially true for the more highly familiar figures. In general, the findings do not support the Piaget & Inhelder hypothesis, but some trends in the data are consistent with a modified form of their hypothesis.

Piaget (1954) and Piaget & Inhelder (1956) have concluded from their extensive studies of the growth of spatial understanding that there is an orderly progression of ontogenetic differentiation of space at both perceptual and imaginal levels. They maintain that at each level the class of topological spatial features is differentiated prior to the class of Euclidean spatial features. One concrete example of this principle is provided by a discrimination task in which young children distinguished topological shape equivalences more readily than Euclidean shape equivalences. Piaget & Inhelder based their conclusions on data derived from a wide variety of tasks and, at least for the tasks they employed, their conclusions account for the data quite well. However, it seemed to us that such a broad principle of the kind offered by Piaget & Inhelder (1956) requires additional research in order to test its generality. Such was the purpose of the present study.

Several researchers have reported on their efforts to verify and extend Piaget & Inhelder's contention that topological spatial differentiation precedes Euclidean differentiation. Page (1959) found that differentiation of the Euclidean features of *rectilinearity–curvilinearity* occurs at least as early as differentiation of the topological features *open–closed* and *hollow–solid*. In a more extensive study with children from 2 years 11 months to 5 years 8 months, Lovell (1959) found that while in general topological features are differentiated earlier than Euclidean features, shapes with curved edges (a Euclidean feature) are as easily identified as are topologically equivalent shapes such as ellipses and polygons. Lovell (1959) argued that it is specific features such as gaps, holes, curves, points, corners, and ins-and-outs, rather than topological or Euclidean features as mutually exclusive classes, which form the basis of differentiation and matching in young children. Pinard & Laurendeau (1966) obtained results similar to those of Lovell, and argued that it may be the perceptual activities involved in figural exploration, rather than the topological or Euclidean features of the stimulus, which are of most importance for discrimination. Other

authors (Zaporozhets, 1965; Braine, 1965; Abravanel, 1968; Vurpillot, 1968) have also reported changes in perceptual activity and scanning mechanisms during ontogenesis, and the importance of such changes for figural differentiation is argued. For example, Braine (1965) has pointed out that a characteristic of the object or figure—its 'focal feature'—determines the initial fixation point for children up to about 5 years, but not for older children. She also noted that scanning proceeds in a downward direction at all ages: downward from the focal point for younger children, and downward from the top of the figure for older children.

From another direction, Munsinger & Kessen (1964) have suggested that a child's discrimination of a visual stimulus is a joint function of stimulus complexity (the number of independent characteristics of the stimulus) and stimulus frequency (also a measure of familiarity with the stimulus figure). This view provides a place for experiential factors in the conceptualization of perceptual development. Since we live in a culture which from an early age stresses the differentiation of objects and shapes in the terms of Euclidean rather than topological properties, it would be expected that any influence of learning would favour the differentiation of Euclidean features.

The purpose of the present study was to examine further the Piaget–Inhelder hypothesis of a developmental shift from topological to Euclidean spatial differentiation by using a new series of figures and a procedure that allows for the direct opposition of topological and Euclidean bases of matching. A second purpose was to investigate the role of figural familiarity in the young child's judgements of relative similarity. The research cited above would suggest that the child may respond differentially to simpler and more familiar figures on the one hand, and to more complex and less familiar figures on the other.

METHOD

Subjects. Subjects were 56 children attending an all-day nursery school in McLean, Virginia. They were white middle-class children, many of whom attended school while their mothers were at work. Ages ranged from 3 years 6 months to 5 years 5 months. Groups 1–4 were composed of subjects with mean ages of 3 years 10 months, 4 years 3 months, 4 years 8 months and 5 years 4 months, respectively. The numbers of subjects in each group were 12, 13, 17 and 14, respectively.

Materials. Two series of figures differing in familiarity and complexity were used. Figures for series A were prepared by cutting shapes out of 3 in. squares of black poster-board and mounting them individually on white 4 × 6 in. file-cards. Within this series, items were designed to be presented in triads, so that each of the two comparison figures of a triad shared Euclidean or topological features with the corresponding standard figure. Fig. 1 presents the figural triads administered to all subjects. The properties of the standard preserved by the comparison stimuli I and II of each triad are as follows: triad 1, (I) *hollowness*, (II) *curvilinearity*; triad 2, (I) *hollowness*, (II) *curvilinearity*; triad 3, (I) *hollowness* and *openness*, (II) *rectilinearity*; triad 4, (I) *hollowness* and *openness*, (II) *hollowness* and *rectilinearity*; triad 5, (I) *hollowness* and *openness*, (II) *hollowness* and *curvilinearity*; triad 6, (I) *continuity*, (II) *rectilinearity*.

Series B repeated many of the conditions of series A, but substituted unfamiliar, rather free forms in place of the more familiar series A forms (see Fig. 1). These forms were cut from 3 in. squares of $\frac{1}{12}$ in. pressed board.* The properties of the standard preserved by the comparison stimuli I and II of each triad are as follows: triads 1 and 2, (I) *rectilinearity*, (II) *solidness*;

* Items for series B were taken from a larger set of figures originally designed to be used as a sorting task. But, as sorting proved too difficult for subjects at these ages, selected items were taken from the set and presented to the subjects in the manner described here.

[illegible] triads 7 and 8, (I) curvilinearity, (II) openness.

[illegible] the comparison stimuli looked 'most like' the standard.

[illegible]

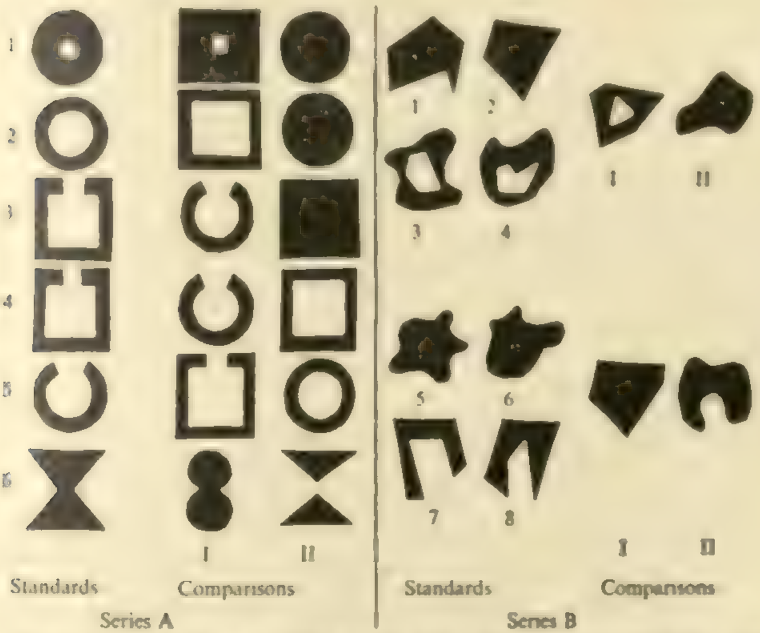


Fig. 1. Triads of series A and B. For series B, triads were formed by presenting each of the eight standards separately with comparisons I and II located to right of given standard in Fig. 1.

Results

Fig. 2 presents graphically the proportion of topologically based similarity judgements by age-group for series A and B. For series A a 4×6 (age-group × triad) analysis of variance yielded no significant main or interaction effects. At all ages the majority of similarity matches was based on Euclidean features. On the other hand, average topological choices (with a maximum possible of 6) by age-group were: group 1 = 1·83, group 2 = 1·62, group 3 = 1·18 and group 4 = 1·64. Inspection of a scattergram revealed no significant age trend in the proportion of subjects consistently high or low in topological choices.

For series B a one-way analysis of variance indicated a trend ($F = 2{\cdot}56$; d.f. = 3, 52; $P < 0{\cdot}10$) from topological to Euclidean bases of matching with age. Yet inspection of the data indicated that for all triads, at all ages, departures from random

matching were in the direction of more frequent Euclidean choices. The average number of topological choices (with a maximum possible of 6) was: group 1 = [illegible], group 2 = 2·40, group 3 = 1·74 and group 4 = 1·32. Inspection of a scattergram suggested a tendency for younger subjects to be more consistent in making topological choices, and for older subjects to be more consistent in making Euclidean choices.

A comparison of average percentage topological choices per age group in series A and B is presented in Fig. 2. A 4 × 2 (age-group × series) analysis of variance suggested an effect of series ($F = 3{\cdot}14$; d.f. = 1, 104; $P < 0{\cdot}10$).

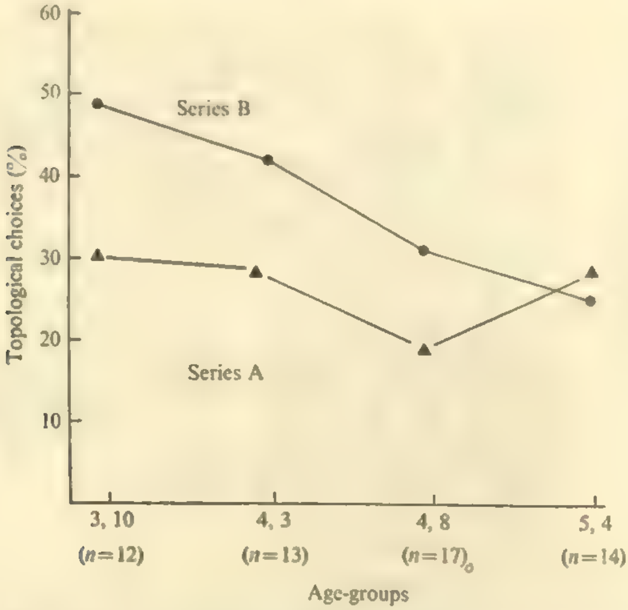


Fig. 2. Proportion of topologically based similarity judgements for series A and B as a function of age.

Discussion

The hypothesis of Piaget & Inhelder (1956), suggesting a progression in spatial development from an ability to differentiate and recognize topological forms of similarity prior to Euclidean forms of similarity, is only partially supported by the present findings. Series B demonstrated a trend of increasing Euclidean-based similarity judgements with age, but series A showed no such trend. In fact, the majority of judgements at all ages was in terms of shared Euclidean features between standard and comparison figures. By questioning the subjects following their participation in the study we confirmed our suspicion that even the youngest children had been introduced to the more common Euclidean shapes during nursery-school life. Further evidence for the greater familiarity of series A was found in subjects' verbalizations during testing. Many subjects spontaneously labelled figures in series A with their correct geometric names, whereas no such labelling was observed for the figures of series B. Thus the figures of series A can be considered as more familiar to

our subjects than those of series B, none of which directly appeared in the school curriculum.

Our findings indicate that the more familiar figures of series A produced the greater number of Euclidean based choices. The importance of familiarity appears at even the youngest ages studied. The mechanism responsible for the observed effects of familiarity (and perhaps figural complexity) cannot yet be specified, but evidence for it appears in other developmental studies (e.g. Munsinger & Kessen, 1964). In conclusion, the Piaget–Inhelder scheme for describing the development of spatial awareness may in the main be correct, but the results of this study do not fully support their position. The development of spatial awareness may proceed from topological to Euclidean bases of differentiation, but factors of familiarity and complexity appear to play a significant role. Familiar figures may incline responding in the direction of Euclidean judgements, even in very young subjects, just as use of familiar or unfamiliar figures and pictures has been shown (Zinchenko *et al.*, 1963) to influence the scanning activity and recognition ability of young children.

References

Abravanel, E. (1968). The development of intersensory patterning with regard to selected spatial dimensions. *Monogr. Soc. Res. Child Dev.* **33** (Whole no. 118.)

Brainerd, L. G. (1965). Age changes in the mode of perceiving geometric forms. *Psychon. Sci.* **2**, 155–156.

Lovell, K. (1959). A follow-up study of some aspects of the work of Piaget and Inhelder on the child's conception of space. *Br. J. educ. Psychol.* **29**, 104–117.

Munsinger, H. & Kessen, W. (1964). Uncertainty, structure, and preference. *Psychol. Monogr.* **78**, no. 9. (Whole no. 586.)

Page, E. I. (1959). Haptic perception: a consideration of one of the investigations of Piaget and Inhelder. *Educ. Rev.* **11**, 115–124.

Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books.

Piaget, J. & Inhelder, B. (1956). *The Child's Conception of Space*. New York: Norton.

Pinard, A. & Laurendeau, M. (1966). The topological nature of the child's early spatial concepts: Piaget's hypotheses revisited. *Int. J. Psychol.* **1**, 243–255.

Vurpillot, E. (1968). The development of scanning strategies and their relation to visual differentiation. *J. exp. child psychol.* **6**, 632–650.

Zaporozhets, A. V. (1965). The development of perception in the preschool child. *Monogr. Soc. Res. Child Dev.* **30**, 82–101.

Zinchenko, V. P., Van Chzhi-Tsin & Tarakanov, V. V. (1963). The formation and development of perceptual activity. *Soviet Psychol. Psychiat.* **2**, 3–12.

(*Manuscript received* 10 *December* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 481–488
Printed in Great Britain

ELATION

By K. T. STRONGMAN, P. E. WOOKEY and R. E. REMINGTON
Department of Psychology, University of Exeter

[illegible] individual procedures.

Like so many other terms that psychologists have inevitably borrowed from every day usage, the word 'elation' can only be as good as its operational definition; otherwise its many connotations generate confusion. Even within academic psychology the word has been used in a variety of ways. From Watson's (1919) Z or love factor, its progress has been roughly dichotomous. Elation has been spoken of in vague terms by some of the neo-Freudians (e.g. Lewin, 1950) and a little more rigorously by some of the neo-behaviourists (e.g. Millenson, 1967). It is the present purpose to concentrate on the latter, but even amongst these there is confusion. However, given a reasonably broad operational definition, it may well be that the concept of elation can be systematized in such a way that more, and perhaps more pertinent, research can be engendered.

Our suggested working definition of elation is that it is a non-instrumental change in the vigour of behaviour which comes about as a result of an upward shift in conditions of reward. Clearly, this is a very broad definition, but broad definitions have the advantage of being capable of further refinement. Before attempting this, however, there are some general points which should be made.

First, elation as defined is an aspect of emotional behaviour and as such is subject to the laws of classical rather than instrumental conditioning. Any change in stimulus conditions can affect behaviour in two ways, i.e. in two directions: backwards in an instrumental sense and forwards in an emotional sense. Thus, apart from any subsequent instrumental effects, an increase in reward may lead to an *immediate*, although perhaps short-lived, respondent change in behaviour. This can, of course, be conditioned classically. The importance of this has already been pointed to in a more general application by Rescorla & Solomon (1967), who conclude 'that *any empirical or theoretical law of Pavlovian conditioning has profound implications for the control of instrumental responding* when the two processes are interactively combined by E's procedures' (p. 170, italics theirs).

It is the present contention that both instrumental and Pavlovian responding are involved in any change in reward conditions, particularly since, in order to assess the immediate emotional effects of a stimulus, it is necessary to have some ongoing instrumental behaviour as a reference point.

Secondly, given that elation has been defined as emotional behaviour which is intimately bound up with instrumental behaviour, the question arises as to whether it has an organizing or disorganizing influence. It is perhaps best to leave this matter until the relevant studies have been reviewed, especially as the final answer is probably equivocal.

Some psychologists (e.g. Millenson, 1967) have argued that to present a non-contingent reinforcer which is either signalled or unsignalled leads to a change in responding which is called elation. Others (e.g. Meyer & McHose, 1968) say that elation results from an increase in contingent reward which leads immediately to greater response vigour than is shown by a control group that always receives the larger reward.

Table 1. *Procedures imposing an upward shift in conditions of reward on established instrumental behaviour*

(The numbers refer to the eight categories of reward increase which are discussed in the text)

Reward increase							
Contingent reward				Non-contingent reward			
Signalled reward		Unsignalled reward		Signalled reward		Unsignalled reward	
Same reward 1	Different reward 2	Same reward 3	Different reward 4	Same reward 5	Different reward 6	Same reward 7	Different reward 8

Given that elation is brought about by an elevation in conditions of reward, then its various forms must depend upon parameters of reward change. An increase (or decrease) in reward can be (i) contingent or non-contingent, (ii) signalled or unsignalled, and (iii) the reward presented under the changed conditions can be the same as or different in quantity or quality from that which went before. These various possibilities are shown in Table 1 which represents a breakdown of the possible methods of changing the conditions of reward. As will be seen when the experimental findings within these various heads are considered in depth, under some conditions response vigour declines and under others it increases. These emotional changes may increase, decrease, or have no effect on the eventual instrumental pay-off. This indicates that a re-examination of the concept of elation would be worth while.

As mentioned previously, in order to quantify elation it is necessary to have established immediately subsequent instrumental behaviour against which emotional effects can be assessed. Hence, only those studies which best fulfil these limiting conditions and which can be subsumed under any of the sections in Table 1 will be discussed.

Category 1: response contingent, signalled, same

This category is logically empty of experiments since the only way of changing the conditions precludes an increase in contingent reward conditions, although the effects of signalling the availability of the normal reward may prove interesting in themselves. Briefly, a possible experiment could involve the establishment of a

[illegible] baseline followed in a test phase by addition of a neutral stimulus for a short period until the time when the normal reward would be available (i.e. somewhat akin to schedules incorporating an 'added clock' (Ferster & Skinner, 1957).

Category 2: response contingent, signalled, different

There is only one relevant study in this category – part of Pliskoff ([illegible]). This will be discussed in Section 6 to which the main body of the experiment is more appropriate.

Category 3: response contingent, unsignalled, same

Empty of experiments. Possible studies would involve a straightforward analysis of occasional unsignalled changes to higher density reinforcement schedules with particular reference to behaviour occurring immediately after the transition point.

Category 4: response contingent, unsignalled, different

Taken at its face value, this section would seem to imply a consideration of the problem of incentive, i.e. differences in instrumental responding under qualitatively and quantitatively different conditions of reinforcement. This has been reviewed elsewhere (e.g. Pubols, 1960) and by virtue of the experimental procedures normally adopted is in fact largely irrelevant to the present discussion. Typically, a rat is allowed to traverse an alley and enter a goal box where some reinforcing stimulus is presented. The subject is then removed from the goal box more or less immediately and returned to its home cage, usually for a 24 hour period. During this time any 'emotion' generated by the reinforcing stimulus would be expected to have dissipated. When the subject is subsequently returned to the experimental situation its behaviour has been studied with respect to two theoretical possibilities: (i) there is no effect of any emotion which reinforcement on the previous trial may have produced, and present responding will be controlled by various absolute factors of the reinforcing stimulus (Capaldi & Lynch, 1967), or (ii) performance will be affected by stimuli which evoke (hypothetical) fractional emotional responses (rg and rf) (Spence, 1960; Amsel, 1958, 1962). But, irrespective of this, few incentive studies have included any measures against which emotional behaviour in the goal may be assessed. For instance, the 'latent learning' studies (e.g. Blodgett, 1929; Tolman & Honzik, 1930*a*, *b*) should provide excellent examples of the direct effect of a reward increase, but their discussion is confined to later effects on performance in the maze.

Only two studies present as much as an anecdotal account of behaviour on presentation of reinforcement. Tinklepaugh (1928), in an atypical study concerned primarily with delayed response discrimination learning in monkeys, found no behavioural changes when a preferred food (banana) was surreptitiously substituted for a less favoured food (lettuce) during the delay, although he felt that 'surprise' must have been present. This is mainly of note in view of the considerable emotional behaviour observed when the monkeys encountered lettuce, having seen the goal being baited with banana. Crespi (1942) also noted that when rats which had not experienced reward shift in the test situation received a relatively small reward, such behaviours as jumping, biting and grooming occurred; whereas, on receipt of a relatively large reward, these behaviours were absent. The 'elation effect' which Crespi

reported is only of nominal interest; although he believed it to be an emotional effect, it was concerned with performance 24 hours after reward shift. Also, more recent studies have cast considerable doubt on the existence of such effects (Spence, 1956).

The main studies which fit the present section all use modifications of the double runway (Amsel & Roussel, 1952). This apparatus allows an evaluation of the direct effects of emotion-producing conditions on subsequent, well-established instrumental behaviour – as well as the possibility of the observation, if not measurement, of the emotional behaviour itself. This is made possible by the manipulation of the reinforcing stimuli in the first goal box (GB1). Despite the suitability of this apparatus for the investigation of reward increase, only four studies have used an 'elation' paradigm, the main emphasis in the literature being on the effects of reward reduction.

Bower (1962) used a triple runway, with rewards of eight, eight and one pellets respectively in the three goal boxes (GB1, 2 and 3). After 198 training trials he found that an increased reward in GB1, of 12 or 16 pellets presented in subsequent test trials on a 50:50 pseudo-random basis, produced a depression of running speed in the second alley (A2) in comparison with the normal reward test trials. However, this study lacked a control group receiving unshifted reward of the greater magnitude.

Two subsequent studies have confirmed Bower's result, although obtaining different effects when performance following a larger reward was compared with that of the pertinent control group. Meyer & McHose (1968) gave 48 acquisition trials with GB1 rewards of one or three pellets and GB2 reward of two pellets, before shifting the GB1 reward to seven pellets on a one in six pseudo-random basis. Running and starting speeds in A2 were lower when following the seven pellet reward than after the training level of reward, but greater than A2 speeds of a control group always receiving seven pellets in GB1. This facilitation of A2 responding in comparison with the controls is called the 'apparent elation effect', although a within-subject depression effect occurred. Strongman & Wookey (1969) trained rats to receive one and two pellets in GB1 and GB2 respectively for 80 trials and then presented two pellets in GB1 on a 50:50 pseudo-random basis for a further 40 trials. As before, faster A2 running speed followed the training reward than the increased reward on test trials. However, in comparison with control groups always receiving either none, two or two, two pellets in GB1 and GB2 there were no significant differences between A2 speeds of the controls and of the experimental group on increased reward trials, whereas the normal reward test trials were characterized by A2 running speeds greater than those found in either control group or in the experimental group on pre-shift trials. The authors interpreted this as a rapidly developing frustration effect.

The other relevant study, Karabenick (1969), used a 2×2 factorial design on reward magnitude in GB1 and GB2. After 30 training trials in which rats received either large or small (15 or one pellets) reward in GB1 and GB2, half the subjects in each group were shifted to the other reward magnitude in GB1 for the full 30 trials in the subsequent test period. Increased reward resulted in decreased A2 running speed, although this still tended to be faster than that of the relevant unshifted controls, lending some support to the 'apparent elation effect'. Within-subject frustration effects were found following decreased reward, but these were not upheld in comparison with unshifted controls. Karabenick points out that to explain this, frustration theory must assume 'comparative frustration', i.e. frustration during

acquisition due to differences in reward in the different goal boxes. This comparative frustration would also be present in any elation studies using different GB1 and GB2 rewards, and tend to obscure any genuine effects. There is at present insufficient evidence to permit any definite conclusions to be drawn on this point, although Wookey & Strongman (1971) have shown that responding during a 10-sec TO between two FR 25s increased during training when rats were never rewarded after the first component but always rewarded after the second, TO responding after the second component remaining near zero. The authors suggested that this is evidence of comparative frustration, since (i) the subject could distinguish between the two parts of the schedule – the response rates were different, (ii) as training continued, subjects developed aggressive behaviours towards the lever during the intercomponent TO, which were not observed at other times, and (iii) the subject never experiences reward in the presence of the stimulus configuration denoting the intercomponent TO, but some of the stimuli present were also part of the configuration denoting reward after the second response sequence. The subsequent effect of the introduction of reward after the first FR on 50 per cent of the trials was to significantly reduce intercomponent TO response rate following reward below that following non-reward, although the former was much greater than that of the baseline rate of a group rewarded after the first FR from the start of training.

Thus these effects of reward increase on 'emotional' responding are similar to the instrumental effects reported by Meyer & McHose (1968), although Wookey & Strongman found no differences in response rate on the FR components following reward and non-reward. It would seem worth while to investigate more rigorously the immediate behaviour following reward (and non-reward) in the first goal box of the double runway, as is suggested in part by Wilton *et al.* (1969).

The most obvious conclusion from these studies is that the investigations within this present category of reward increase have been unsystematic. Studies involving differently sized mazes, different numbers of training trials, different amounts of reinforcement and different schedules of its delivery are unlikely to yield consistent results. However, it does seem possible to draw two tentative conclusions. (i) An increase in reward in the first goal box of a double runway after a period of training at a lower level of reward, leads to a within-subjects performance decrement in the second alley. (ii) Under some conditions the performance does not fall as low as that of unshifted controls always receiving the larger reward. The within-subjects depression may be the effect of other, confounded factors such as drive reduction with the larger rewards. It has been suggested that in the range 0–4 pellets reward in GB1, running speed is an inverse non-monotonic function of reward level (Meyer & McHose, 1968; Daly, 1968). However, even response suppression may be less simple than this. When zero reward in GB1 has been used, running speed has sometimes been slower than that of rewarded subjects during acquisition (Wagner, 1959), or not different (Strongman & Wookey, 1969).

Categories 5 *and* 6: *non-contingent, signalled, same and different*

Several operant studies of elation fall into these two categories of change in reward structure, i.e. non-contingent, but signalled presentation of a similar or different free reward. Studies of this kind have used Estes & Skinner's (1941) design

for evoking the conditioned emotional response. They have, however, substituted positive for negative reinforcement as the unconditioned stimulus (US) and generalized to using both positive and negative baselines on which the classical paradigm has been superimposed.

Taken together, conditioned responding to a conditioned stimulus (CS) signalling either positive reinforcement or the removal of negative reinforcement, represents that part of Millenson's (1967) model of emotional behaviour subsumed under the general heading of elation. Millenson's model allows that a CS signalling either presentation of S+ or removal of S− produces the same elation state. This is in contrast to the two different states produced by signalled presentation of S− (fear) or removal of S+ (anger). In effect, presentations of an S+ or S− are convenient to arrange within the Estes–Skinner paradigm, whilst their removal is difficult. This has led to removal of the opportunity to respond for an S+ (time-out; TO), rather than the S+ itself, being used to produce what is called anger (Herrnstein, 1955; Leitenberg, 1965; Azrin *et al.*, 1966). The logical converse would be removal of the necessity to respond to avoid an S−, whilst to maintain consistency, Millenson's model should use the signalled removal of continuing aversive stimulation. As far as the present authors are aware, no studies of either kind have been carried out, although as will be shown below a negative baseline has been used.

Herrnstein & Morse (1957) first used signalled free-food reinforcement to produce an elational effect. Pigeons were trained to respond on DRL 5′ (this typically leads to very low rates of responding). After training, illumination of the key with a yellow light for 2 min. (CS) at random intervals, produced no noticeable effects on response rate. The usual 5 sec. food-hopper presentation was then programmed to occur after 1 min. of the CS. This procedure produced sudden and marked overall increases in response rate, an effect which, over continuing sessions, was replaced by a differentially high rate of responding during the CS. In all cases, the rate was greatly in excess of that which was reinforced on a DRL schedule. Herrnstein & Morse discount the possibility that the initial elation effect is the result of superstitious response-reinforcement contingencies, whilst allowing that the high rate produced by free reinforcement could be superstitiously *maintained*. Their contention is that the defining property of a reinforcer is not its sole behavioural consequence and suggest the use of different reinforcers and schedules with the same paradigm would serve to broaden the generality of the effect.

Azrin & Hake (1969) have adopted this suggestion, but with contradictory results. They ran 18 rats on variable interval (VI) 1 min. schedules of positive reinforcement. During the baseline phase a 10 sec. CS (either a clicker or a blinking light), unpaired with any CS, produced no change in rate of responding. During testing, a US was made contingent on the end of CS, which occurred at random in all phases. A return to baseline phase followed testing. Using food or water as contingent reinforcers, and food, water or intercranial stimulation (ICS) as the US, the design was such that for each animal the reinforcer was either quantitatively and/or qualitatively different from the US. The results showed that a suppression of responding occurred in all but three subjects. By programming the US such that it did not occur within 1 sec. of a bar press, two of the subjects which had shown response facilitation demonstrated a marked suppression after 23 sessions. Interestingly, these two subjects were in the

group using water both as reinforcers and US. All these changes in rate were extinguished in the return to baseline phase where the US was again omitted.

There are a number of differences between Herrnstein & Morse's and Azrin & Hake's studies which perhaps account for their apparently divergent results. (i) The former used pigeons, the latter rats; species differences have already been established within the CER literature (e.g. Leitenberg, 1965). (ii) The reinforcement density generated by DRL schedules is typically much lower than that generated by VI schedules. (iii) In Azrin & Hake's study, even if deprivational variables were matched between the subjects working for food and water, then depriving them of the reinforcer to be used as a US on the following day would have worked to lower their motivational state according to the food/water interaction discussed by Bolles (1967). The nature of this interaction was unclear. (iv) Herrnstein & Morse used a 2 min. CS, presenting the US after 1 min.; Azrin & Hake used a 10 sec. CS with the US contingent on its offset. The latter is an unusually short CS interval, which may have only sampled a subject's initial, orientating responses.

Henton & Brady (1970) have confirmed that CS duration is an important variable in this context. Running rhesus monkeys on a DRL 30-sec. schedule, they showed an elational acceleration greater than 150 per cent with an 80 sec. CS, but no effect with shorter CS durations.

Notwithstanding the above, Azrin & Hake's study does show positive conditioned suppression to be a general effect, at least during a short CS. Two studies by Pliskoff increase the generality of this finding. In the first (1961) he used pigeons on a multiple VI VI schedule which alternated every 15 min. The switch was signalled during the last 3 min. of each component by warning stimuli, During a baseline both VIs were of the same mean interval (mult VI1 VI1 for one group, mult VI10 VI10 for the other). During testing, all four birds were changed to mult VI10 VI1. Thus the change to a different reward structure was non-contingent but signalled. All subjects showed a downward shift in response rate during the CS which signalled the more favourable schedule. An acceleration was seen when the less favourable schedule was signalled. Clearly, for interval responding these rate changes failed to increase pay-off in any way. In a subsidiary part of the same study, Pliskoff programmed fixed interval (FI) 3 to run concurrently with the CS, thus in the present terms making the change in reward structure both signalled and contingent (category 2). Here the signalled change from VI10 to VI1 produced scalloping of the record, whereas the opposite change produced the non-contingent US pattern in one subject and a 'somewhat irregular' performance in the other. Pliskoff indicates that the high to low reinforcement schedule transition produces effects similar to those produced by signalled TO (Herrnstein, 1955; Ferster, 1958), i.e. an acceleration, so the effects of the low to high transition may be expected to be suppressive. These results show clear support for Azrin & Hake's data.

In further refinement, Pliskoff (1963) incorporated a baseline control rate (VI) over which signalled changes to higher or lower reinforcement schedules (both VI) were superimposed at random. This modification allowed the same reinforcement potential during each CS, and upheld the result of the earlier study; again signalled higher reinforcement schedule led to suppression.

Can the data from these three studies be meaningfully combined using a unitary

elational concept? Azrin & Hake assert that their findings suggest that the elation paradigm produces 'emotionality' (i.e. the complex of respondents which characterize discharge of the autonomic nervous system), and emotionality depresses response rate. Pliskoff's data support their case in that signalling the high pay-off 'arm' of a mult VI VI produces response suppression. However, Pliskoff (1963) also reports that signalled low pay-off VI or extinction increases the rate. This result reflects those of Leitenberg (1965) using TO as a US. Furthermore, studies which have superimposed a CS for positive reinforcement, but conditioned out of the experimental situation, have all reported facilitatory effects (Estes, 1948; Morse & Skinner, 1958; Bower & Grusec, 1964). If unsignalled *secondary* positive reinforcement leads to increased response rates, the effect should be conditionable to a CS, albeit that CS duration may need to be greater than 10 sec.

Data from negative conditioned suppression studies may help in analysing the relative merits of the experiments mentioned so far in this section. For example, Blackman (1968*b*) has shown that the rate of extinction of conditioned suppression is a function of schedule-dependent response rate (holding constant reinforcements acquired), and a function of reinforcement density (when rate is constant). Blackman (1968*a*) has also shown that for some low levels of shock (< 1 mA), bar-pressing during the CS on a DRL schedule shows an increase in probability, though for higher shock levels suppression occurs. Clearly, Azrin & Hake's and Herrnstein & Morse's experiments differ along dimensions of both schedule and rate. From non-quantitative observations of subjects during CS on DRL, Blackman reports that timing responses (cf. Laties *et al*., 1965) are disrupted, and this may produce a decrease or an increase in rate. This allows the interpretation that any strong positive or negative stimulus disrupts (rather than organizes) behaviour to be applied to Herrnstein & Morse's results. If a key factor in DRL responding is the timing response, then this may be disrupted by a CS signalling S+, and lead to an increased frequency of criterial responding. This we can assume to be a high probability response. Any such response will be subject to adventitious contingencies. From this one would predict that facilitation should occur to a greater extent if a subject has more than one response with a high probability, and could be tested by observing behaviour in which more than one criterial response occurs.

The obvious importance of the gross nature of the schedule used in 'elational conditioning' is well illustrated by Coulson & Walsh (1968). They reversed the conditions of the original Estes–Skinner study by using a negative baseline (VI 20 avoidance SS 1) and presenting 10 sucrose pellets as a US, after a 1 min. CS. The authors adopted a Rescorla baseline (cf. Rescorla, 1967) and during test paired CS and US incorporating a delay circuit to avoid adventitious reinforcement. With rats at 70 per cent of *ad lib*. weight, the CS produced response facilitation, which was maintained when various response-CS delays or no delays were incorporated. Coulson & Walsh claim that this data will fit Sheffield's (1966) theory that when a consummatory response is signalled but unable to occur, an animal responds with 'general excitement' and that this excitement is 'channelled' into the response that is being performed. However, Azrin would presumably predict a suppression.

In fact, Coulson & Walsh's result may be artifactual in that subjects were required not to respond for 10 sec. in order to produce the US. On a VI shock schedule, not

responding for this 10 sec. could lead to shock. A subject would therefore be reinforced for increasing rate during a CS by the avoidance contingency. As with all CER studies, the effect could be superstitiously maintained during the no-delay period.

From these studies only one firm conclusion can be drawn, i.e. presenting a CS signalling free positive reinforcement will reliably produce a change in response rate during that CS. The nature of this change may be to suppress or to facilitate criterial responding, but as yet no one explanation can account for all the data. Millenson & DeVilliers (1971) have recently proposed that conditioned anxiety could be re-examined in terms of the motivational properties of the stimulus conditions. Applying this model *in toto* to CER using a positive US should yield a prediction of CS facilitation for all positive baselines. As has been shown, this is not necessarily the case. Overall, some disruption would seem to occur in some kind of responding, but this may in fact lead to the facilitation of a criterial response.

Category 7: *non-contingent, unsignalled, same*

Empty of experiments. A possible experiment could involve the presentation of a non-contingent, unsignalled reward during a period of low response density on any reinforcement schedule. Appropriate schedules would be extended FI and DRL.

Category 8: *non-contingent, unsignalled, different*

Unsignalled increases in overall pay-off structure commonly produce a brief and transient facilitation in responding. Azrin (1960) trained rats on a VI 1 schedule, and after a baseline had been obtained, punished with shock every response that subjects made whilst keeping the reinforcement contingencies in operation. Finally, the punishment contingency was removed. The initial effect of shock was to decrease response probability, but after 20 sessions the animals were again responding at baseline rates. When punishment was discontinued, a sudden, temporary response facilitation was observed. This transient increase in response rate when an unsignalled, non-contingent change in reward occurs is usually termed the punishment contrast effect. Holz *et al.* (1963) have found the same effect on a DRL schedule where punishment raised rather than lowered the frequency of reinforcement, by suppressing responding.

Thus response facilitation occurs after a period when reinforcements have been decreased or increased by punishment. Further, the effect is produced by a direct manipulation of the reinforcement contingencies. Ferster & Skinner (1957) have shown that on a schedule where intermittent reinforcement is interrupted by periods of TO, a transient increase in rate is observed following the offset of the TO period. Similarly, Reynolds (1961) trained pigeons to respond on a mult VI 3′ VI 3′, the duration of each component being signalled by a red or green light. When the rates in red and green were equal, the schedule was changed to mult VI 3′ red : Ext green. The procedure caused the 'green' rate to extinguish in the normal way, whilst an increase in response rate was observed in red. This effect is, of course, that of Ferster & Skinner. However, when the original contingencies were re-instituted, a transient increase in response rate above the initial baseline was observed in green, with a concomitant decrease in red. This is behaviour contrast.

Clearly, Estes's (1944) theory that the organism is 'making up' for lost reinforce-

ments after punishment is unable to account for the punishment contrast findings since no (or few) reinforcements were lost. Nevertheless, the effects are all transient and emotional – or unnecessary – in nature. Also, Pliskoff (1961) has shown that they are all conditionable to a signal preceding a change in reinforcement structure.

Discussion

As regards the effects of the various experimental procedures that we have gathered together under the heading of elation, all the sources of evidence add up to equivocation. An upward shift in reward does have immediate emotional effects on ongoing instrumental behaviour and some paradigms have been developed for studying its conditioned effects. However, there does not seem to be any consistency in these behaviour changes; sometimes there is an enhancement or facilitation of responding, sometimes there is an impairment or suppression. So, although the procedures under discussion have a common root in reward increase, their effects are dissimilar. Further, when the procedures lead to a suppression of ongoing behaviour, they are still not difficult to relate to the everyday notion of elation, which has connotations of excitation and euphoria, which may be facilitatory or disruptive.

The present conclusions lead to some unfortunately all too common possibilities for the future. The term elation could be restricted to procedures and not effects. It would be clearly inappropriate to tie it to effects and not procedures. For example, no one would describe Amsel's double-runway rats as elated when they find no reward in GB1. It is only acceptable to speak of effects when they are reasonably clear-cut. Again, for example, it is appropriate and convenient to use the phrase 'frustration effect' for the well-established and reliable increase in response vigour found when an animal experiences non-reward where it had commonly experienced reward. With respect to elation viewed in this way, the effects are simply not that clear.

To carry this argument further, it may be preferable not to use the term to cover a large number of procedures, but rather to concentrate on clarifying the effects of the individual procedures themselves. This conclusion, as ever, calls for extensive parametric investigations. The present review and analysis may be of some help in considering and designing such investigations.

The preparation of this paper was supported by grant G 969/269/B from the Medical Research Council to the first author.

References

Amsel, A. (1958). The role of frustrative non-reward in non-continuous reward situations. *Psychol. Bull.* **55**, 102–119.

Amsel, A. (1962). Frustrative non-reward in partial reinforcement and discrimination learning: some recent history and a theoretical extension. *Psychol. Rev.* **69**, 306–328.

Amsel, A. & Roussel, J. (1952). Motivational properties of frustration. I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.* **43**, 363–368.

Azrin, N. H. (1960). Sequential effects of punishment. *Science, N.Y.* **131**, 605–606.

Azrin, N. H. & Hake, D. F. (1969). Positive conditioned suppression: conditioned suppression using positive reinforcers as the unconditioned stimuli. *J. exp. Analysis Behav.* **12**, 167–173.

Azrin, N. H., Hutchinson, R. R. & Hake, D. F. (1966). Extinction-induced aggression. *J. exp. Analysis Behav.* **9**, 191–204.

Blackman, D. (1968*a*). Conditioned suppression or facilitation as a function of the behavioral baseline. *J. exp. Analysis Behav.* **11**, 53–61.

Blackman, D. (1968*b*). Response rate, reinforcement frequency and conditioned suppression. *J. exp. Analysis Behav.* **11**, 503–516.

Blodgett, H. C. (1929). The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.* **4**, 113–134.

Bolles, R. C. (1967). *Theory of Motivation.* New York: Harper & Row.

Bower, G. H. (1962). The influence of graded reductions in reward and prior frustrating events upon the magnitude of the frustration effect. *J. comp. physiol. Psychol.* **55**, 582–587.

Bower, G. H. & Grusec, T. (1964). Effect of prior Pavlovian discrimination training upon learning an operant discrimination. *J. exp. Analysis Behav.* **7**, 401–404.

Capaldi, E. J. & Lynch, D. (1967). Repeated shifts in reward magnitude: evidence in favor of an associational and absolute (noncontextual) interpretation. *J. exp. Psychol.* **75**, 226–235.

Coulson, G. & Walsh, M. (1968). Facilitation of avoidance responding in white rats during a stimulus preceding food. *Psychol. Rep.* **22**, 1277–1284.

Crespi, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *Am. J. Psychol.* **55**, 467–517.

Daly, H. B. (1968). Excitatory and inhibitory effects of complete and incomplete reward reduction in the double runway. *J. exp. Psychol.* **76**, 430–438.

Estes, W. K. (1944). An experimental study of punishment. *Psychol. Monogr.* **57**, no. 3. (Whole no. 263.)

Estes, W. K. (1948). Discrimination conditioning. II. Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *J. exp. Psychol.* **38**, 173–177.

Estes, W. K. & Skinner, B. F. (1941). Some quantitative properties of anxiety. *J. exp. Psychol.* **29**, 390–400.

Ferster, C. B. (1958). Control of behavior in chimpanzees and pigeons by time out from positive reinforcement. *Psychol. Monogr.* **72**, no. 8. (Whole no. 461.)

Ferster, C. B. & Skinner, B. F. (1957). *Schedules of Reinforcement.* New York: Appleton–Century–Crofts.

Henton, W. W. & Brady, J. V. (1970). Operant acceleration during a pre-reward stimulus. *J. exp. Analysis Behav.* **13**, 205–209.

Herrnstein, R. J. (1955). Behavioral consequences of the removal of a discriminative stimulus associated with variable-interval reinforcement. (Unpublished doctoral dissertation, Harvard University.)

Herrnstein, R. J. & Morse, W. H. (1957). Some effects of response-independent positive reinforcement on maintained operant behavior. *J. comp. physiol. Psychol.* **50**, 461–467.

Holz, W. C., Azrin, W. H. & Ulrich, R. E. (1963). Punishment of temporally spaced responding. *J. exp. Analysis Behav.* **6**, 115–122.

Karabenick, S. A. (1969). Effects of reward increase and reduction in the double runway. *J. exp. Psychol.* **82**, 79–87.

Laties, V. G., Weiss, B., Clark, R. L. & Reynolds, M. D. (1965). Overt 'mediating' behavior during temporally spaced responding. *J. exp. Analysis Behav.* **8**, 107–116.

Leitenberg, H. (1965). Is time-out from positive reinforcement an aversive event: a review of the experimental evidence. *Psychol. Bull.* **64**, 428–441.

Leitenberg, H. (1966). Conditioned acceleration and conditioned suppression in pigeons. *J. exp. Analysis Behav.* **9**, 205–212.

Lewin, B. D. (1950). *The Psychoanalysis of Elation.* New York: Norton.

Meyer, P. A. & McHose, J. H. (1968). Facilitative effects of reward increase: an apparent 'elation effect'. *Psychon. Sci.* **13**, 165–166.

Millenson, J. R. (1967). *Principles of Behavioral Analysis.* New York: Macmillan.

Millenson, J. R. & DeVilliers, P. A. (1971). Motivational properties of conditioned anxiety. In R. M. Gilbert & J. R. Millenson (eds.), *Reinforcement Processes: Schedule-induced and Schedule-dependent Effects.* London: Academic Press. (In press.)

Morse, W. H. & Skinner, B. F. (1958). Some factors involved in the stimulus control of operant behavior. *J. exp. Analysis Behav.* **1**, 103–107.

Pliskoff, S. (1961). Rate-change effects during a pre-schedule-change stimulus. *J. exp. Analysis Behav.* **4**, 383–386.

Pliskoff, S. (1963). Rate change effects with equal potential reinforcements during the 'warning' stimulus. *J. exp. Analysis Behav.* **6**, 557–562.

Pubols, B. H. (1960). Incentive magnitude, learning and performance in animals. *Psychol. Bull.* **57**, 89–115.

Rescorla, R. (1967). Pavlovian conditioning and its proper control procedures. *Psychol. Rev.* **74**, 71–80.

Rescorla, R. A. & Solomon, R. L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* **74**, 151–182.

Reynolds, G. S. (1961). Behavioral contrast. *J. exp. Analysis Behav.* **4**, 57–71.

Sheffield, F. D. (1966). New evidence on the drive induction theory of reinforcement. In R. N. Haber (ed.), *Current Research in Motivation*. New York: Holt, Rinehart & Winston.

Spence, K. W. (1956). *Behavior Theory and Conditioning*. New Haven: Yale University Press.

Spence, K. W. (1960). *Behavior Theory and Learning*. Englewood Cliffs, N. J.: Prentice-Hall.

Strongman, K. T. & Wookey, P. E. (1969). Frustrative reward. *J. exp. Psychol.* **82**, 183–184.

Tinklepaugh, O. L. (1928). An experimental study of representative factors in monkeys. *J. comp. Psychol.* **8**, 197–236.

Tolman, E. C. & Honzik, C. H. (1930*a*). Degrees of hunger, reward and non-reward, and maze learning in rats. *Univ. Calif. Publ. Psychol.* **4**, 241–256.

Tolman, E. C. & Honzik, C. H. (1930*b*). Introduction and removal of reward, and maze performance in rats. *Univ. Calif. Publ. Psychol.* **4**, 257–275.

Wagner, A. R. (1959). Role of reinforcement and non-reinforcement in an 'apparent frustration effect'. *J. exp. Psychol.* **57**, 130–136.

Watson, J. B. (1919). *Psychology from the Standpoint of a Behaviorist*. New York: Lippincott.

Wilton, R. N., Strongman, K. T. & Nerenberg, A. (1969). Some effects of frustration in a free responding operant situation. *Q. J. exp. Psychol.* **21**, 367–380.

Wookey, P. E. & Strongman, K. T. (1971). Emotional and instrumental effects of reward shift. *Psychol. Rec.* **21**, 181–189.

(*Manuscript received* 29 *October* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 493–500
Printed in Great Britain

ANISEIKONIA: THE ACHILLES HEEL OF THE CLASSICAL THEORY OF VISION

BY JAMES G. TAYLOR*

There are two kinds of naturally occurring aniseikonia. Parallactic aniseikonia (retinal disparity) is one-dimensional, and it is commonly supposed to be a major factor in the genesis of the perception of depth. Asymmetric aniseikonia occurs when ocular convergence is asymmetric, so that a fixated object is nearer to one eye than to the other, and it is two-dimensional. It has fatal implications for any theory based on the assumption that visual perception is fully determined by events in the visual system, from the retina to the striate area, and consequently is seldom mentioned in the literature. It has no such implications for the behavioural theory of perception (Taylor, 1962), which fully explains two perceptual constancies that persist through a wide range of variation in aniseikonia. The behavioural theory offers a simple explanation of the delay in the development of perceptual distortion when aniseikonia is produced artificially by means of a lens.

If a lens that magnifies along the horizontal but not the vertical axis is placed in front of one eye, there is at first no disturbance of binocular vision. Distortion develops gradually and reaches a maximum within a few minutes. This phenomenon, which has not been found with any other kind of transformation that has been applied to the ocular input, has attracted much less attention than it deserves. Seagrim (1967*a*, *b*, 1969) has tried, with various experimental techniques, to find a reason for the delay, but has failed. The reason for his failure will become evident when we take account of the aniseikonias that occur naturally. But what is more important is that it will become clear that the phenomenon is fatal to the classical theory of vision that has prevailed, with only minor modifications, since the time of Helmholtz.

What I mean by the classical theory of vision is any theory that is based on the assumption that the structure of visual perception is fully determined by the optical properties of the eyes and the morphology of the visual receptor system from the retina to the striate area and adjoining cortical structures, but makes no reference to the behaviour that is initiated by stimulation of this system. Variations on this theme are irrelevant to the following discussion, and no reference need be made to any of them. It is the central theme itself that I am proposing to attack.

There are two kinds of natural aniseikonia. The first type is associated with binocular parallax. If a rectangle, standing in a vertical plane, is inclined at an oblique angle to the median plane of the observer's body, its width subtends unequal visual angles at the two eyes, but its height subtends equal angles, i.e. parallactic aniseikonia is one-dimensional. According to standard doctrine this aniseikonia is one of the cues for the perception of depth, and indeed it has sometimes been described as the primary cue.

In addition to this there is a two-dimensional aniseikonia that occurs whenever an object is viewed with asymmetric convergence, and its amount can be easily calculated. Suppose, for example, that a subject with an interocular distance of 70 mm

* 110 Freedom Avenue, Riverside, Kitwe, Zambia.

turns his eyes so that his line of vision is at an angle of 45° with the median plane of his head, and views a small object at a distance of 50 cm from a point midway between the nodes of his eyes, the visual angle subtended by the object at the near eye is approximately 9·4 per cent larger than the angle subtended at the far eye. The degree of aniseikonia is a negatively accelerated decreasing function of the distance of the object when the direction of vision is held constant. Thus if the distance is decreased to 25 cm aniseikonia increases to 22 per cent. If the distance of the object is held constant, the degree of aniseikonia is an increasing function of the angle between the median plane of the head and a line joining the object to the cyclopean eye. If aniseikonia is defined as the percentage by which the visual angle at the right eye exceeds the angle at the left eye, it has a positive value when the median plane of the head passes to the left of the object, falls to zero when that plane intersects the object and becomes increasingly negative as the plane veers to the right.

Let us now specify that the object occupies a fixed position relative to a stationary observer, that its visible surface is a plane and that this plane is orthogonal to a line joining its centre to the cyclopean eye. The specification of orthogonality is intended to reduce parallactic aniseikonia to zero, or at least to the negligible proportions entailed by changes in the position of the cyclopean eye when the head is rotated.

If the observer fixates the object while rotating his head about its vertical axis, the visual angles subtended by the object at the two eyes change continuously in opposite directions, so that aniseikonia varies through a range of positive and negative values. Now if this aniseikonia plays the same kind of role in perception that is commonly attributed to parallactic aniseikonia, we should expect two things to happen during this exercise. First, because of the horizontal retinal disparity the object should be perceived as rotating about its vertical axis, reaching orthogonality with the line of vision when the median plane of the head intersects it. And secondly, the varying retinal disparity in the vertical direction should give rise to a diplopia that becomes more marked as the plane of the head deviates in either direction from the object. But in fact nothing of the kind happens. The perceived orientation of the object remains invariant and there is no trace of diplopia. If two horizontal lines on slightly different levels are presented in a stereoscope (which of course demands symmetric convergence), there is, for each subject, a minimum angular separation of the lines above which it is impossible to fuse them into a single perceived line. Yet fusion occurs regularly when greater degrees of vertical disparity are produced by asymmetric convergence.

The head-rotating exercise has revealed two perceptual constancies that have hitherto been neglected, possibly because of their embarrassing implications. In both of them stimuli whose function has long been assumed to be understood produce results that are incompatible with standard doctrine. Hence a different set of principles will be required to account for them.

I have already (Taylor, 1962) expounded a general theory of perceptual constancy, but it will be useful to spell it out in detail as it applies to the present case.

Since the constancies operate at all possible positions and orientations of the object, it will be sufficient to explain their origins in relation to an object in one fixed position, and then the theory can be generalized to all other positions. Let us assume that the object is at shoulder height, that it is within reach of the hands, that its

visible surface is inclined at an angle θ to the median plane of the body and is orthogonal with a line from its centre to the cyclopean eye. The first constancy determines that the perception of the angle θ is invariant, the second that the object retains its singularity in perception.

Let there be n different directions of the head, arrived at by rotation about the vertical axis alone. No additional principles are involved if the head is allowed to rotate about its other two axes as well, but the proposed restriction helps to simplify the exposition. In each of the n positions the eyes may face in any one of many directions, but I shall admit only one, namely that which fixates the centre of the object. (I shall justify this drastic pruning after showing the derivation of the constancies.) Then there are n patterns of binocular retinal stimulation, all produced by the same object and all centred at the foveas, differing from each other only with respect to the degree of aniseikonia. It is clear that we cannot find the source of either perceptual invariance if we restrict our search to the visual afferent system. But associated with each of the n directions of the head there is a distinctive pattern of proprioceptive afference, originating in the muscle, tendon and joint receptors of the neck, and in the receptors located in the extrinsic muscles of the eyes.

Since, for the given position of the object, both the degree of aniseikonia and the pattern of proprioceptive afference are determined by the direction of the head, we have n unique patterns of stimulation, each having a complex visual component, v, and a complex proprioceptive component, p. The component v could be produced by an object in any one of several other positions, but in that case the component p would have a different form. The uniqueness of the pattern (v, p) arises from the fact that it can be generated only by an object in the specified position and at exactly one of the n positions of the head. We then have n patterns of the form (v, p), and these patterns have to be shown to be equivalent to one another with respect to the perceived angle θ, and with respect to the singularity of the object. The relation of equivalence may be symbolized by

$$E(\theta) \equiv [(v, p); (v, p) \rightarrow \text{perception of } \theta].$$

In words, the equivalence class of theta is defined as a set with elements of the form (v, p), such that each (v, p) gives rise to perception of the same angle θ. The source of the equivalence is not to be found in the elements themselves, since they all differ from one another. Our problem then is to find a mechanism that will generate equivalence from this collection of unequal elements.

The answer turns out to be remarkably simple, although the mechanism itself is anything but simple. Suppose the subject wants to operate on the object in some way, such as writing or drawing on it, or painting, cleaning or polishing it, or even stroking it with a finger to discover its texture. In all such operations success depends on the hand moving in the plane of the object, or in a plane parallel to it. If, while any element of $E(\theta)$ is in operation, the hand moves in another plane, it loses contact or collides with the object, and that response is extinguished. Only movements in the right plane are reinforced. Since each of the n elements is unique, it can be conditioned, and is in fact conditioned, to the same response. It is this invariance of the response that makes the elements equivalent to one another, and not their intrinsic character.

For if the r components were changed by some transformation applied to the ocular input, the equivalence would be destroyed at first, but would soon be restored if the subject got plenty of opportunity to correct his behavioural errors.

The essence of the behavioural theory of perception is that perception is a state of multiple simultaneous activity in the neural engrams mediating learned responses to patterns of stimulation determined jointly by stimulus energies from the environment and by the proprioceptors. Or more briefly, perception is a state of multiple simultaneous efferent readiness. This implies that the content of perception is determined not by the intrinsic character of the afferent processes but by those properties of the environment that serve to reinforce successful responses. Thus in the case we have been considering the reinforcing agency depends primarily on the angle θ, and consequently what is perceived is precisely that angle. The varying degrees of aniseikonia are in themselves completely irrelevant. They are merely components of the elements of $E(\theta)$ that serve to trigger the appropriate response, or the state of readiness for it. Like the symbols of the British postal code they serve, in conjunction with other components of the afferent pattern, to specify an address to which the response is directed.

The term 'multiple simultaneous readiness' implies that many other events are taking place at the same time as those I have been describing, but for an account of these the reader will have to consult *The Behavioral Basis of Perception* (Taylor, 1962). There is, however, one other set of events that must be described here, namely that which is responsible for our second constancy, the invariant singularity in the vertical aspect of the object despite a wide range of variation in vertical retinal disparity.

The behavioural theory suggests that the perception of size is related to behaviour adapted to the actual size of the object and has nothing to do with the size of the retinal image. The visual angle subtended by an object can vary through two orders of magnitude without affecting perceived size. As we have seen, the n elements of $E(\theta)$ are all unique, and therefore each of them is capable of being conditioned not only to a response that is exactly adapted to the plane of the object but also to a response that is exactly adapted to its size, such as placing fingers and thumb at just the right distance apart to grasp the object. Regardless of what the retinal images say, the object itself has only one size, and therefore there is only one separation of fingers and thumb that will suffice to grasp it. This is the only response that gets reinforced, and therefore it is the only response that gets conditioned to all the elements of $E(\theta)$. It is a singular response and not a pair of unequal responses that somehow get averaged. Hence the behavioural theory enables us to calculate that naturally occurring aniseikonic stimulation generates singular perception.

I can now justify my apparently arbitrary decision to admit only one direction of the eyes for each of the n positions of the head in defining the equivalence class $E(\theta)$. Each of the m different directions in which the eyes can face involves a unique pattern of afferent events, including visual and proprioceptive components, but since the stimulus object that is responsible for the visual components is assumed to be invariant in position, size, shape and orientation, it is evident that the responses conditioned to any one of the m patterns must be conditioned to each and all of the others. That is to say, the m patterns themselves constitute an equivalence class, and

therefore one of them, selected on an arbitrary basis, will suffice to represent them all. The equivalence class $E(\theta)$ has, in fact, not n but nm elements.

Let us return now to parallactic aniseikonia. Psychologists who accept the classical doctrine concerning the genesis of sensory and perceptual experience are in general satisfied that they have a fairly good understanding of the function of this kind of aniseikonia. This understanding is in fact illusory, because their calculations are based on the assumption, seldom explicitly stated, that convergence is symmetrical. The relation between retinal disparity and perceived depth may be correctly formulated for this limiting case, but it is certainly not true of the general case in which convergence is asymmetrical. And even if it were true that in the limiting case the striate area and adjacent structures could process the input from the retina to generate three-dimensional perception, the rules for processing would necessarily give rise to unstable perception in the general case.

To illustrate the difficulty let us return to the equivalence class $E(\theta)$, which refers to an object whose surface is orthogonal to the line of vision. If we rotate the object so that its angle with the median plane of the body is θ_1 instead of θ, it will now be oblique to the line of vision, so that there is an element of parallactic aniseikonia in addition to asymmetric aniseikonia. This parallactic component is not a constant, however, since the angle between the two lines of regard decreases as asymmetry of convergence increases, and of course it is subject to alteration by the horizontal component of the asymmetric aniseikonia.

If we now construct an equivalence class $E(\theta_1)$, we can see that each of its elements is different from all the others, as in $E(\theta)$. If we compare element i of $E(\theta_1)$ with element i of $E(\theta)$, it will be seen that the p components are identical, unless eye movements involving convergence on the vertical edges of the object are admitted. The v components differ with respect to their horizontal dimensions, and in a different way with respect to their vertical dimensions, since the vertical edges are no longer equidistant from either eye. It follows that each element of $E(\theta_1)$ differs not only from all the other elements of its own class but also from all the elements of $E(\theta)$; and this is true no matter what value we assign to θ_1. It follows also that each of the elements of $E(\theta_1)$ can be conditioned to a response exactly adapted to the angle θ_1, thereby generating constancy in the perception of that angle.

Now if we isolate the horizontal disparity components from the elements of $E(\theta)$ and $E(\theta_1)$, it can easily be seen that a proportion of the disparities occurring in $E(\theta)$ are also to be found in elements of $E(\theta_1)$; and this has a direct bearing on the classical theory that there is a functional relation between the degree of horizontal retinal disparity and the perceived angle of slant. It has sometimes been assumed, or implied, that this relation is at the same time a relation of cause and effect, in the sense that the perception is directly generated in the striate area by the cortical events resulting from the disparity. A clear example of this is Metzger's (1953) theory of binocular perception of depth, which leans heavily on Gestalt principles but retains the basic ingredients of classical theory. To test the theory it is first of all necessary to have a clear understanding of what a functional relation is.

A function is most conveniently defined in the terminology of set theory. If A is a set with elements a, B is a set with elements b, and F is a set with elements (a, b), then F is a function if and only if each a that occurs in F is paired with exactly one b.

On the other hand, each b may be paired with more than one a. To test the classical theory, let A be the set of horizontal retinal disparities arising in the situation defined above, and B the set of perceived angles. We construct F by assembling all the (a, b) pairs that can occur in the actual situation. Since many of the horizontal disparities incorporated in the elements of $E(\theta)$ occur also in elements of $E(\theta_1)$, this means that any given a may be paired with more than one b, and therefore that F is not a function. In that case there can be no relation of cause and effect between A and B.

We next define C as a set with elements $c = (v, p)$, i.e. the set that is partitioned by the equivalence classes $E(\theta)$; a set D whose elements are the angles of the planes in which the hand moves in response to elements of C; and a set F_1 with elements (c, d), including only pairs that occur in practice. It is evident that F_1 is a function, since each c is paired with exactly one d, although each d is paired with all the elements of an equivalence class of C. If we now assemble a set F_2 with elements (c, b), where B is the set of perceived angles, it is clear that this too is a function, since each pattern of stimulation is paired with exactly one perceived angle.

Finally we assemble, from F_1 and F_2, a set F_3 with elements (d, b), such that for each (d, b) b is the perceived angle that is associated with the response d. F_3 is not only a function but a one-to-one correspondence, for not only is each d paired with exactly one b, but each b is paired with exactly one d. As this correspondence cannot be fortuitous, we can confidently assert that the mechanism that is responsible for the assembly of F_1 is also responsible for the assembly of F_2. The mechanism for F_1 is learning, and thus we arrive again at the behavioural theory of perception.

We are now ready to consider artificially induced aniseikonia and the problem that has baffled Seagrim. The essence of the matter is that the pattern of stimulation is one that never occurs in natural circumstances. The lens produces a one-dimensional aniseikonia, but the pattern of stimulation does not include the differences between the visual angles subtended by the vertical edges of the test object that are found when this kind of aniseikonia occurs naturally. In Expt. VII (Seagrim, 1969) the test object was placed at an oblique angle to the median plane of the body, thereby producing a pattern compounded of natural and artificial aniseikonia. But again it is a pattern that cannot occur in nature. The subject's head is immobilized, and the test object is in the middle of the visual field, so that there is no two-dimensional aniseikonia. But even if asymmetric convergence were allowed, not one of the patterns of stimulation would be one that the subject had encountered before.

Here then is the simple key to the solution of the problem. The patterns of stimulation are all new, and there has been no opportunity to condition them to responses appropriate to the orientation of the test object. The state of readiness generated by the environment contains no elements determined by the unnatural aniseikonia, and therefore no new property of perception is detected or reported.

In this respect the aniseikonic lens differs from other optical devices that have been used to transform the retinal input. For example, a wedge prism has the effect that the v component of a pattern of stimulation relating to a given object is paired with a p component with which, in normal vision, it would be paired only if the object occupied a different position in space. But the pairing is not an impossible one; it has occurred again and again, and a suitable response has been conditioned to it. Reverting to the analogy of the British postal code, we can say that the effect of

wedge prisms is to alter one of the symbols in the code while leaving the others unchanged, but in such a manner that the new arrangement of symbols represents a real address. The aniseikonic lens, on the other hand, alters the symbols in such a way that no existing address is specified. The code word is not in the dictionary.

If aniseikonia is produced binocularly, one lens magnifying horizontally and the other vertically, each of the distorted pictures is intelligible in itself, presenting no patterns of stimulation that do not evoke a state of readiness for response, but again the binocular patterns are 'impossible' and can contribute nothing to perception. In these circumstances behavioural readiness is likely to be controlled by one eye. It may alternate between the eyes, or part of the field may be controlled by one eye and part by the other, as often happens in retinal rivalry. If only one lens is used, the field of the unencumbered eye will have the preference, particularly if it is the dominant eye. In any case perception remains normal for a time. There may be some distortion of either the horizontal or the vertical dimension, but none in the dimension of depth.

The phenomenon of retinal rivalry is not irrelevant to our problem. When red and blue squares are presented in a stereoscope the colours do not mix, not because they cannot mix but because an object cannot be simultaneously both blue and red. The red-naming and blue-naming responses are in practice mutually exclusive. If there are stimuli tending to evoke both and to locate them in one and the same place, one of them must be inhibited, since there is no learned response to the total binocular pattern.

Similarly, it is impossible for an object to be simultaneously both 25 and 26 in. wide. The artificially induced aniseikonic pattern tends to evoke responses adapted to both widths, and as these responses are addressed to one and the same object, one of them has to be inhibited. On the other hand, an object 30 in. wide can be so placed that it subtends at the two eyes separately the horizontal visual angles that would be subtended at both eyes by objects of 25 and 26 in. standing in a frontal parallel plane. In that case the total pattern of binocular stimulation addresses a single response to the object, and it is reasonably accurately adjusted both to the true width of the object and to its orientation.

We must now ask why the artificial aniseikonic pattern, having failed to evoke this response initially, ultimately does evoke it, so that normal perception gradually gives place to distorted perception. The answer, I believe, is that the evocation of readiness for two incompatible responses, even in alternation, directed to the same object, is ultimately intolerable. Real solid objects are singular, and the organic system demands a singular response and searches for the most likely one. It tries to find a real address for the hitherto unknown code message. Since the alien feature of the message is a persistent one-dimensional aniseikonia, it is not surprising that this should play a dominant part in the selection of a singular response. But because the conditions of the experiment allow the subjects to look at the object but not to handle it, the new singular response is not an overt one but merely a state of readiness for it. And that means that the object is perceived as having changed its shape and orientation.

If the subjects were required to wear the lens not just for a few minutes but for some days continuously, so that the states of readiness referred to above could be trans lated into overt actions, these would of course suffer extinction, and new responses

would be conditioned to the altered patterns of binocular stimulation, which would thereby become codes for real instead of non-existent behavioural addresses. And as this process of re-education continued, the perceived world would gradually return to its true shape.

The principles I have invoked here are quite general, and are not confined to the case I have discussed. For example, they can be used to explain the neglect of two-dimensional aniseikonia in the literature of vision. It is not as if the facts were difficult to discover. Anyone with a knowledge of plane trigonometry can find them, and it is reasonably certain that they have been discovered again and again. But the trouble is that there is no satisfactory response that will fit them into the framework of the classical theory of vision. Any tentative explanatory response that is compatible with the basic principles of that theory will be extinguished by the absurdities that are revealed. But if the facts cannot evoke any response that can be reinforced, the final outcome is the same as if the facts did not exist. There is no awareness of them.

Despite this functional blindness the facts do exist, and if there is no answer to them within the framework of classical theory, then that theory has got to go.

REFERENCES

METZGER, W. (1953). *Gesetze des Sehens*, 2nd ed. Frankfurt am Main: Kramer.

SEAGRIM, G. N. (1967*a*). Stereoscopic vision and aniseikonic lenses. I. *Br. J. Psychol.* **58**, 337–350.

SEAGRIM, G. N. (1967*b*). Stereoscopic vision and aniseikonic lenses. II. *Br. J. Psychol.* **58**, 351–356.

SEAGRIM, G. N. (1969). Stereoscopic vision and aniseikonic lenses: further studies in latency. *Br. J. Psychol.* **60**, 307–313.

TAYLOR, J. G. (1962). *The Behavioral Basis of Perception*. New Haven: Yale University Press.

(*Manuscript received* 25 *August* 1969; *revised manuscript received* 7 *December* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 501–504
Printed in Great Britain

ON THE TERMS 'SUBLIMINAL PERCEPTION' AND 'SUBCEPTION'

By W. J. REES

Department of Philosophy, University of Leeds

The terms 'subliminal perception' and 'subception', as now employed in psychology, are widely agreed to be unclear. It is here argued that it is necessary to distinguish between a developmental and an analytical use of these terms and, more importantly, to recognize that on the analytical use, which is also the prevailing use, the term 'subliminal perception' does not designate a form of perception but a form of epiperception. It is further argued that, in the light of these considerations, it is possible to disperse at least some of the confusions now attending the use of these terms.

'Subliminal perception' normally designates stimulus-response (SR) processes in which the *stimuli* are below the level of awareness, but processes also by which human beings somehow respond to some features of their environment. 'Subception' normally designates SR processes which are similar except that, in addition, the *responses* are confined to the autonomic nervous system (or, as some have used the term, to some system other than that underlying verbal report). Briefly, and without too much oversimplification, subliminal perception might be defined as *discrimination without awareness*, whereas subception might be defined as *autonomic discrimination without awareness*.

Such definitions are frequently given and they seem technically competent and scientifically innocent. Yet, in fact, the resulting concepts are unclear and volatile in meaning. The root of the trouble lies in a basic obscurity about the nature of the distinction between subliminal and supraliminal perception. This distinction can be viewed in two ways. It can be treated as a developmental distinction or as an analytical distinction. As a developmental distinction it has always been treated as a morphogenetic distinction. As an analytical distinction there is always a distinct temptation to treat it as a distinction between two species of the same genus, whereas, upon investigation, it turns out to be a distinction of another kind, the nature of which is left entirely obscure. In order to elucidate these matters, let us consider the types of processes which may be labelled 'subliminal' and the types of processes which, in contrast, are labelled 'supraliminal'.

By 'subliminal perception' one may mean, and is often meant, a perceptual process which is preconscious in a morphogenetic sense. A subliminal perception, in this case, is that *stage* in the development of a perceptual process which immediately precedes the point at which the process becomes a conscious one. A supraliminal perception is the remaining and conscious stage of that process. Subliminal perceptions, so understood, are more or less identical with Leibniz's *petites perceptions* (see Gerhardt, 1882), and identical with what some have now called the 'microgenetic' processes in perception (Werner, 1956). They are also identical with many of the processes which others have classified under the heading of 'perceptual defence'. Dember (1960) writes: 'A conception of the perceptual process as one which operates

in stages is not in itself very difficult to accept. Nor is the idea unacceptable that events which influence behaviour can occur without the individual's awareness.

The distinction between subliminal and supraliminal perception is usually regarded, however, as an analytical and not as a developmental distinction. Subliminal and supraliminal perceptions are then not stages within processes but distinct processes. We are also invited to suppose that they are two species of one genus, that is to say, two types of processes having certain essential features in common and differing only in so far as the processes are below awareness in the one case and above awareness in the other. But here we encounter difficulties. If a supraliminal perception is not a stage within a perceptual process but a complete process, it can only be what we ordinarily refer to as a perceptual process *simpliciter*, i.e. the same process as that which includes subliminal and supraliminal stages when regarded developmentally. In that case also a subliminal perception, by contrast, can only be a process which is not perceptual at all, assuming a consistent use of the word 'perceptual'. Subliminal perception is a class of processes which resemble in some ways the class of processes which we call perception, but it exists *alongside* perception and is *not a subclass* of it. To put it briefly, subliminal perception is not a species of perception; it is more appropriately described as a species of *epiperception*. Correspondingly, supraliminal perception is also not a species of perception; it is nothing but perception under an assumed name.

The term 'subception', as normally employed, designates a species of subliminal perception in this sense. It is also accordingly a species or, more accurately, a subspecies of epiperception.

These are important matters which have implications over a wide range of theoretical issues. Leaving aside those issues which are of interest mainly to philosophers, such as the question whether or not the notion of subliminal perception is self-contradictory and the question whether or not the existence of subliminal perception would imply the existence of subconscious mind, there are two matters which are of more direct importance to psychology.

In the first place, it is possible to dispose of the problem whether or not subliminal and supraliminal perception lie on continua of SR processes. This question is no longer one question but two different questions and it so happens that each requires a different answer. (*a*) If by 'subliminal and supraliminal perception' we understand preconscious and conscious perception in the *morphogenetic* sense, then these do lie on a continuum, but *not* on a continuum of SR *processes*. They constitute adjacent stages in growth or development, but the growth or development takes place within the single SR process and not within a series of such processes. (*b*) If by 'subliminal and supraliminal perception' we understand epiperception and perception, then there is no reason to suppose that these lie on a continuum of any kind. Even though there may be continuities *within* both perception and epiperception, there is no reason to suppose that either of these is continuous *with* the other. If there were such reasons, the distinction which we are here considering would be a developmental and not an analytical distinction. In fact, and in the light of the evidence, any assumption of continuity can arise only from confusing the two distinctions.

Secondly, it is possible to recover and preserve scientific distinctions which are vital to psychology but which are in danger of being obliterated by present

terminology. If I am right, the use of the term 'subliminal perception' is vitiated by two tendencies which mutually support each other. There is a tendency to confuse or conflate two senses of 'subliminal perception', and there is a tendency to treat subliminal and supraliminal perception as two species of a common genus. However different these two tendencies may appear to be—the one linguistic and the other 'scientific'—they are, in fact, logically complementary. If we conflate the two senses of 'subliminal perception' this is tantamount to forming a class of processes which have one characteristic in common, namely that they are below awareness—this, despite the fact that what we are bundling together are complete processes (epiperceptions) in the one case and a stage within other processes (preconscious perceptions) in the other case. This encourages us to distinguish this class from another class having the same characteristics except that the processes are now above awareness—this, despite the fact that what we are now bundling together are complete processes (perceptions) in the one case and a stage (the conscious stage) within these same processes in the other case. Hence we have an egregious distinction which looks like a distinction between two species of the same genus. But looking at the matter in another way, if we begin by supposing that the distinction is one between species of a common genus, we can find differentiae for the species of subliminal perceptions only in the subliminal nature of the stimuli, which in turn necessitates a lumping together of two different things, the one a process of a certain kind and the other a stage within a process of another kind, which is finally tantamount to conflating two different notions of subliminal perception. Surely, this is a vicious circle of confusions. It is, however, a case where scientific confusion is bound up with linguistic confusion. To remove the one, it is essential, and indeed sufficient, to remove the other. If I am right, we now see how both may be removed.

An analogy may help to clarify some of the points I have made. Let us take raw recruits and soldiers on a stage. A raw recruit represents a stage in the military training of an individual not unlike the preconscious stage of perception when perception is regarded morphogenetically. A soldier on a stage has, in a broad way, about the same relationship to a real soldier as epiperception has to perception: each resembles the other in certain respects, but the one is not a species of the other. Now let us put these 'soldiers' together in a single class and call them submilitary soldiers so as to distinguish them from trained soldiers of military age. This is roughly what has happened to the notion of subliminal perception. Just as the notion of submilitary soldiers obliterates distinctions which are militarily important, so the notion of subliminal perception obliterates distinctions which are psychologically important. What is necessary in both cases is to return to more basic distinctions and not to make play with classifications which are superficial and of little explanatory value. For 'submilitary soldiers' we must rewrite 'raw recruits' or 'soldiers on the stage'. For 'subliminal perception' we must rewrite 'preconscious perception' or 'epiperception'. The conceptual problems surrounding the notion of subliminal perception are then, if not at an end, at least reduced to more manageable proportions.

I am indebted to Dr N. F. Dixon of University College London for drawing my attention to this problem and for his comments on an earlier draft of this paper. I have also profited from his own published papers on the topic under discussion (Dixon, 1958, 1964, 1968) and from the first four chapters of his book *Subliminal Perception: the Nature of a Controversy*, which was in press at the time of writing. The responsibility for any deficiencies in the paper is, of course, mine.

References

Dember, W. N. (1960). *The Psychology of Perception.* New York and London: Holt, Rinehart & Winston.

Dixon, N. F. (1958). The effect of subliminal stimulation upon autonomic and verbal behavior. *J. abnorm. soc. Psychol.* **57**, 29–36.

Dixon, N. F. (1964). Communication without awareness: the implications of subliminal perception. *J. psychosom. Res.* **8**, 337–341.

Dixon, N. F. (1968). Perception without awareness: a reply to K. M. Banreti-Fuchs. *Acta psychol.* **28**, 171–180.

Dixon, N. F. (1971). *Subliminal Perception: the Nature of a Controversy.* London: McGraw-Hill.

Gerhardt, C. I. (1882). *Die philosophischen Schriften von Gottfried Wilhelm Leibniz*, vol. 5. Berlin. (English translation: *New Essays Concerning Human Understanding by Gottfried Wilhelm Leibniz.* Chicago and London, 1917.)

Werner, H. (1956). Microgenesis and aphasia. *J. abnorm. soc. Psychol.* **52**, 347–353.

(*Manuscript received* 19 *November* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 505–511
Printed in Great Britain

PRAGMATIC ASPECTS OF NEGATION IN SENTENCE EVALUATION AND COMPLETION TASKS

By E. R. CORNISH

Department of Psychology, North-East London Polytechnic

The appropriateness of a sentence such as 'The circle is not all red' was examined in relation to circles in which the proportion of red was varied. Expt. I showed that subjects took longest to evaluate this sentence when the circle contained only a small proportion of red. Expt. II showed that the tendency to complete an incomplete sentence with 'red', and to consider such a completion as appropriate, increased with the proportion of red in the circle. These results suggest that the appropriateness of a negative copulative sentence is frequently a function of a correlated affirmative presupposition.

Recent experimental studies have focused on the function of negative sentences in normal discourse. Wason (1965) has provided evidence to suggest that their function 'is generally to emphasize that a fact is contrary to an expectation', and Greene (1970*a*, *b*) has shown that negative copulative sentences serve to signal a change in meaning. These views imply that a presupposition is often involved in this type of sentence for speakers and listeners. Greene (1970*b*) claims that 'the negative carries out an *active* meaning change process on a prior affirmative statement. An affirmative, on the other hand, implies nothing about its converse'. The presupposition for a negative copulative sentence consists of what is denied, i.e. it is the same sentence with 'not' deleted; for 'The circle is not all red' it would be 'The circle is all red'.

If negatives are used to deny a presupposition, then the following hypothesis is plausible. The appropriateness of a negative depends upon the extent to which the situation resembles the state of affairs described by the presupposition. Of course, if the presupposition is a completely correct description of the situation, the negative will be false and totally inappropriate. But given that the negative is, in fact, true, then the closer the situation is to the presupposition, the more appropriate it will be. Hence it was predicted that the statement 'The circle is not all red' would be increasingly appropriate as the proportion of red in the circle increased.

This prediction was tested by measuring the speed and accuracy with which such sentences were evaluated in Expt. I, and by the way in which sentence frames were completed in Expt. II.

Experiment I

Method

The task was to evaluate the truth of the sentence 'The circle is not all *y*' (where *y* was 'red', 'blue', 'green' or 'yellow') in relation to a circle divided into two areas of different colour (red, blue, green or yellow). There were four stimulus conditions, depending on the proportion of the circle occupied by the colour mentioned in the sentence. These can be described by taking the sentence 'The circle is not all red' as an example. The stimulus conditions for this sentence are shown in Fig. 1. The circle for condition 11 was 11/12 red and 1/12 another colour; for condition 6 it was 6/12 red and 6/12 another colour; for condition 1 it was 1/12 red and 11/12 another colour; for condition 0 it was divided equally into two colours, neither of which was red.

Subjects. Forty-eight first-year psychology undergraduates from University College London were tested individually.

Apparatus. The sentences and circles (appearing 2·75 in. in diameter) were presented on cards by a projection tachistoscope, to which a Dekatron was attached. Three keys were situated in front of the screen. The central one switched on the visual display and started the timer; either of the outer ones stopped the timer; they were labelled 'true' and 'false' respectively, the labels being switched between subjects on a random basis.

Procedure. The subjects were allotted at random to one of the four conditions, with 12 subjects in each condition. Each subject had one test trial, to which the correct answer was always 'true'. This was preceded by four practice trials, which consisted of the presentation of an affirmative sentence of the form 'The circle is all *y*' and a homogeneous circle. The correct answer to trials 1 and 4 was 'true' and to trials 2 and 3 'false'.

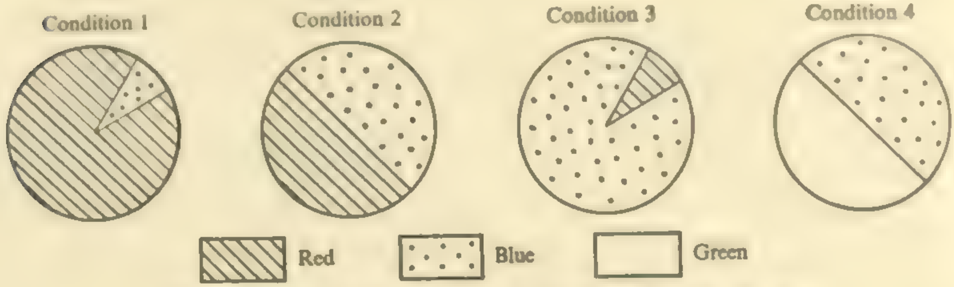


Fig. 1. Examples of the test stimuli in the four conditions for the sentence 'The circle is not all red'.

The four colours appearing in the stimuli and the four colour names appearing in the sentences were randomized with the restriction that each appeared once during the first four trials.

The orientation of the stimulus was randomly varied in four different positions.

The subject was seated in front of the apparatus. He was told that on each trial he would be presented with a sentence about the colour of a circle. He was asked to read the sentence aloud and to press down the central key when he was sure he could remember it. This response would remove the sentence from view and present the circle. He was then to press down either the 'true' or the 'false' key, whichever was appropriate. His response was to be as fast as possible, but accuracy was more important than speed. The trials proceeded without interruption, with an interval of 7 sec. between the removal of the sentence and the presentation of the corresponding stimulus.

Results

The mean response times on the test trial in each condition are plotted in Fig. 2. An analysis of variance on the data (transformed so as to yield homogeneity of variance) showed that there was a reliable difference between the conditions ($P < 0{\cdot}01$) and a significant departure from linearity ($P < 0{\cdot}01$), allowing for the fact that the conditions were not equally spaced along the independent variable. It is evident from the graph that this was due to condition 1; and this was confirmed by Duncan's multiple range test ($P < 0{\cdot}05$).

Thus, although a significant conditions effect was obtained, the order of difficulty was not as predicted. The most surprising result was the exceptional difficulty of condition 1 (where 1/12 of the circle was in the predicated colour), which even exceeded that of condition 0 (where none of the circle was in the predicated colour).

Experiment II

In the second experiment the same problem was investigated using a productive task. The subjects were required to complete sentences so that they corresponded with stimuli similar to those used in Expt. I.

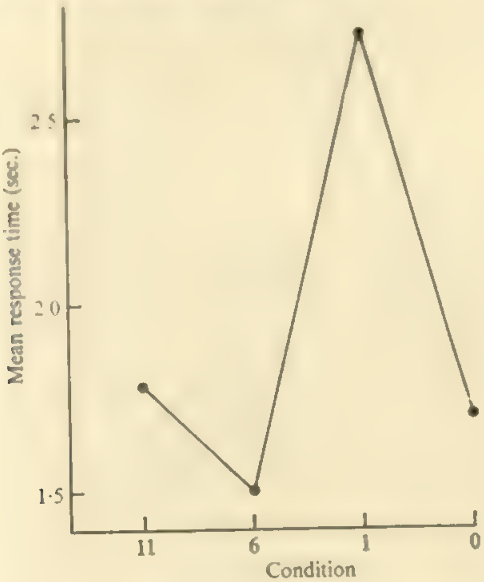


Fig. 2. Mean response times on the test trial.

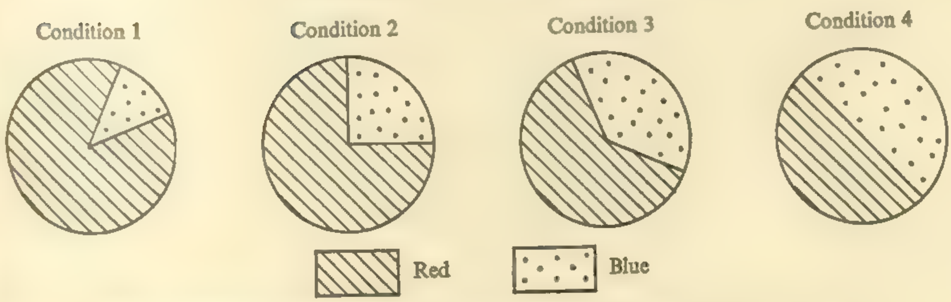


Fig. 3. Examples of the test stimuli in the four conditions for one colour set.

Method

The incomplete test sentence was 'The circle is not all...', the possible responses being 'red' and 'blue'. Examples of the stimuli are shown in Fig. 3. There were four conditions. In condition 1 the circle was 7/8 red and 1/8 blue; in condition 2 it was 3/4 red and 1/4 blue; in condition 3 it was 5/8 red and 3/8 blue; and in condition 4 it was half red and half blue. The orientation of the sectors was held constant. For half of the stimuli the colours were reversed. Subjects had two trials on each condition, one on each colour arrangement. The trials were presented in a random order.

There were two groups: in group Se the sentence preceded the stimulus; in group St the stimulus preceded the sentence.

Subsequently, the subjects were given a completed sentence and asked to rank the four test stimuli in order of appropriateness.

It was predicted that the sentence would be completed in terms of the colour of the larger sector in the circle, i.e. subjects would tend to respond 'red' when the majority of the circle was red. Such responses were called 'appropriate' (A); responses in terms of the smaller sector were called 'inappropriate' (I). Since this criterion cannot be applied in the control condition 4, the responses were classified on the basis of position, i.e. in terms of the top right sector and the bottom left sector. The more frequent of these two types of response was labelled 'A' and the less frequent 'I', to go against the prediction that there would be a chance distribution of 'A' and 'I' responses in condition 4, but an increase in the number of A responses over conditions 3 to 1. It was also predicted that response times would be faster for A than for I responses, and that the speed of A responses would increase through conditions 4 to 1. Finally, it was predicted that, when the test stimuli were ranked for their appropriateness to a completed sentence, they would be placed in the following increasing order of appropriateness: 4, 3, 2, 1.

Subjects. A further 48 non-psychology undergraduates from University College London were individually tested.

Apparatus. The apparatus and materials were similar to those used in Expt. I with appropriately modified stimuli. There were two keys, labelled 'red' and 'blue' respectively; the position of the 'red' key, on the left or the right, was randomly determined for each subject.

Procedure. The subjects were allotted at random to one of the two groups, with 24 subjects in each group.

The eight test trials were preceded by four practice trials, which consisted of the presentation of an incomplete, affirmative sentence, 'The circle is all...', and a homogeneous red or blue circle. These four trials involved two red and two blue circles in a randomized order.

The subjects in group Se were told that they would be presented with an incomplete sentence concerning the colour of a circle. They were to repeat the sentence aloud and to remember it, as it would then be removed from view. The circle would then be presented and their task was to press whichever of the two keys they thought was appropriate to complete the sentence. They were warned never to press both keys, and that their reaction would be timed.

The subjects in group St were told that they would be presented with a circle. They were asked to describe it in their own words and to remember it, as it would then be removed from view. An incomplete sentence describing the circle would then be presented and their task was to press whichever of the two keys they thought appropriate to complete the sentence.

In the ranking task which followed, four test stimuli were presented, one from each condition but all with the same arrangement of colours, which was randomly determined for each subject. The subjects were asked to rank the stimuli in the order of their appropriateness to the given sentence.

Table 1. *Frequencies of appropriate and inappropriate responses*

Condition ...	1		2		3		4	
Group ...	Se	St	Se	St	Se	St	Se	St
Response pattern AA	17	17	13	15	12	15	9	5
AI	3	6	9	8	9	8	11	15
II	4	1	2	1	3	1	4	4

Results

Responses. The responses were classified as 'A' or 'I' as described under 'Method'. Since each subject had two trials on each condition, three response patterns were possible: AA (appropriate responses on both trials), AI or II. The relative frequencies of these patterns are shown in Table 1 for all conditions and groups.

The prediction that there would be more A than I responses in conditions 1–3 but a chance distribution in condition 4 was confirmed by binomial tests on the number of subjects with AA as opposed to II response patterns (AI response patterns do not affect the issue). Significant values ($P < 0{\cdot}02$ in all cases) were obtained for conditions 1, 2 and 3 and non-significant values for condition 4.

The prediction that there would be an increase in the number of A responses over conditions 4–1 was confirmed by Jonckheere trend tests (Group Se: $\tau = 0{\cdot}34$, $P < 0{\cdot}05$; group St: $\tau = 0{\cdot}39$, $P < 0{\cdot}01$).

Response times. The mean response times associated with A and I responses (based on a mean in the case of AA and II response patterns) are shown in Fig. 4. In the majority of cases the times are in the predicted direction, with A responses having a shorter latency than I responses. However, related t tests on the differences within

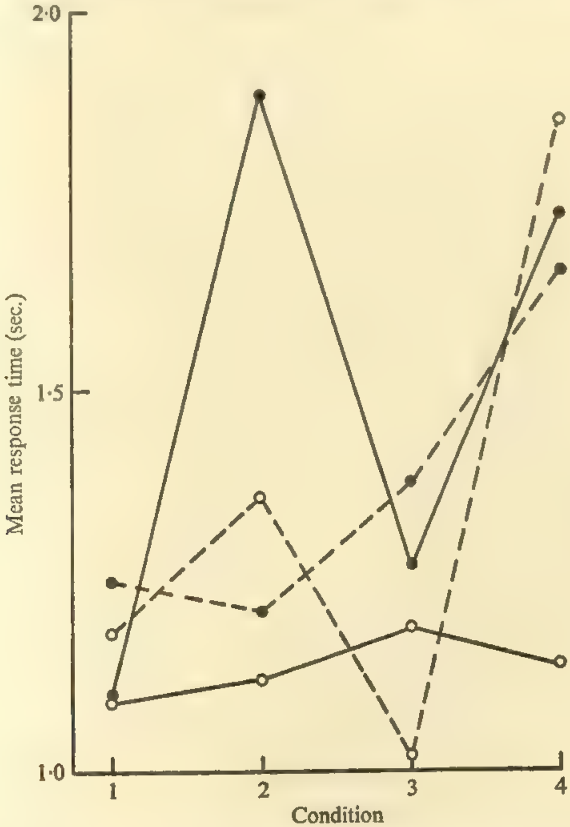


Fig. 4. Mean response times for appropriate and inappropriate responses. Group Se: ●——●, A responses; ●- - -●, I responses. Group St: ○——○, A responses; ○- - -○, I responses.

each condition were not significant. These tests were based solely on the AI response patterns in order to hold constant any effect of individual differences between subjects.

The predicted increase in response times to A responses through conditions 1–4 was tested by Jonckheere combined trend tests (on means in the case of AA response patterns). The prediction was confirmed for group Se ($\tau = 0{\cdot}60$, $m = 20$, $P < 0{\cdot}04$) but not for group St ($\tau = 0{\cdot}37$, $m = 22$, $P > 0{\cdot}05$).

Ranking. In group Se, 17 subjects ranked the stimuli in the predicted order, five ranked them in the reverse of the predicted order, and two in other orders. In group St, 14 subjects ranked them in the predicted order, seven in the reverse

of the predicted order, and three in other orders. Assuming all *observed* orders to be equally likely, the frequency of rankings in the predicted order is highly significant in both groups ($P < 0.001$ on a binomial test).

Discussion

The results of both experiments provide considerable support for the hypothesis that the sentence 'X is not all y' applies with increasing appropriateness as X increases in y. Subjects in Expt. II certainly completed sentence frames in accord with this principle; and the latencies of their appropriate responses also followed the predicted pattern, at least in group Sc where the sentence was presented prior to the stimulus. A similar confirmation was also provided by the way in which subjects evaluated the appropriateness of the stimuli to a given sentence.

However, there were also some unexpected results. The predicted linear trend in response times was not obtained in Expt. I. Instead the mean response time for condition 1, where only 1/12 of the circle was in the predicated colour, was significantly longer than those for the other three conditions, including condition 0 in which the stimulus contained none of the predicated colour. This suggests that emphasis should be placed not on the facilitation which occurs when 'The circle is not all red' is applied to a circle which is predominantly red, but on the difficulty in applying the sentence to a circle which contains only a minimal area of red. There are two factors which may help to explain this finding. First, the sentence is ambiguous. 'X is not all y' can mean either that 'X is predominantly y' or that 'X is not y at all'. It is possible that the interpretation of the sentence 'The circle is not all red' gives rise to two expectations: the 'pragmatic' and more probable expectation of a circle which is predominantly red, or the 'formal' and less probable expectation of a circle which is not red at all. In conditions 11 and 6 the sentence could be verified by reference to the 'pragmatic' interpretation, and in condition 0 it could be verified by reference to the 'formal' interpretation. This may be why condition 0 was easier than expected. The stimulus in condition 1, however, does not resemble either interpretation very closely. It does not resemble the 'formal' interpretation, which demands a complete absence of red, in the way that the stimulus in condition 6 resembles the 'pragmatic' interpretation.

A second contributing factor may be the perceptual simplicity of the equally divided circles in conditions 6 and 0 which could make them easier to evaluate.

It might be argued that an alternative explanation of the difference between conditions 11 and 1 is that it is the relative size of the colour mentioned in the sentence which determines the ease of evaluation. If this were the case, the same factor should influence the ease of evaluating affirmative sentences. However, subsequent experiments involving the evaluation of affirmative and negative sentences referring to smaller or to larger areas of colour did not support this.

The ambiguity of the test sentence was also supported by the results of the ranking test. Here, in addition to the larger number of rankings in the predicted order, there were a surprisingly large number of rankings in exactly the opposite order. Presumably, these rankings occurred when the 'formal' interpretation took precedence over the 'pragmatic' interpretation. This is evident in the following typical introspection:

'Condition 4 is the most appropriate for "The circle is not all blue" because it is the least blue.'

It might be supposed that subjects who adopted the 'formal', as opposed to the 'pragmatic', interpretation for the ranking task would have adopted a similar approach in the completion task. However, there was no evidence that 'formal' subjects made more I responses or were slower in making A responses than 'pragmatic' subjects, although a larger number of 'formal' subjects needs to be examined before any firm conclusion can be drawn.

Coding. Expt. II highlights the problem of coding, a familiar phenomenon in this type of task where a series of repeated measures is required for a stable result, but where intelligent subjects can rarely resist anticipatory guesses. They may try to 'jump the gun' or to develop special strategies for dealing with the situation. Such strategies are intrinsically interesting, though they may be peculiar to the particular experimental situation. In this connexion, a comparison of the two groups is revealing. There is reason to suppose that in group St both the sentence and the stimulus were available to the subject for the development of strategies, whereas in group Se only the stimulus was available. This is because the items that came second for group St, i.e. the sentences, were identical during the test trials. Although the subjects did not know at first what sentences would appear, they probably were eventually able to predict them with a fair degree of accuracy and confidence. This would enable them to select their response prior to the presentation of the sentence, and with only a small probability of having to change it. Such a strategy was not possible for group Se, since the items which came second, i.e. the stimuli, were not identical. This distinction probably explains two differences between the performance of the groups. The significance levels for the difference between the frequencies of AA and II response patterns were higher and more consistent in group St than in group Se; and the trend in response times to A responses was significant for group Se but not for group St. Thus, subjects in group St reached a higher level of 'appropriate' responding than the subjects in group Se.

There is a need for the development of less artificial tasks. An example of a freer situation used by the author is the colouring in of a figure to correspond with a given sentence. It is hoped that studies employing this technique will be the subject of a future paper.

This paper is based on part of an unpublished University of London Ph.D. thesis. I am indebted to Dr P. C. Wason who supervised the research, of which this forms part, and to Dr A. R. Jonckheere and Mr J. D. Valentine for statistical advice. I am also grateful to the referees and to Dr P. N. Johnson-Laird for helpful criticisms of an earlier version of this paper.

References

Greene, J. M. (1970*a*). The semantic function of negatives and passives. *Br. J. Psychol.* **61**, 17–22.

Greene, J. M. (1970*b*). Syntactic form and semantic function. *Q. J. exp. Psychol.* **22**, 14–27.

Wason, P. C. (1965). The contexts of plausible denial. *J. verb. Learn. verb. Behav.* **4**, 7–11.

(*Manuscript received* 18 *February* 1970; *revised manuscript received* 8 *September* 1970)

Br. J. Psychol. (1971), **62**, 4, *pp.* 513–517
Printed in Great Britain

AN EVALUATION OF ELICITATION PROCEDURES FOR PERSONAL CONSTRUCTS

By FRANZ R. EPTING, DAVID I. SUCHMAN and CARL J. NICKESON

Department of Psychology, University of Florida, Gainesville, Fla., U.S.A.

The opposite method (OM) and the difference method (DM) for eliciting personal constructs were evaluated with regard to the number of bipolar constructs produced. This evaluation is an investigation of how the constructs were used in a grid procedure and how they were reported in a structured interview. Comparing the two methods, using the grid procedure, the OM produced a greater number of bipolar constructs. Responses to the interview were incomplete. Explanations for this observed difference between methods are proposed.

Within the purview of Personal Construct Theory (Kelly, 1955) an individual orders his world by placing interpretations upon it. These interpretations, termed personal constructs, are bipolar abstractions of encountered events with the two poles standing in contrast to each other. Focused primarily on interpersonal relationships and personal characteristics, constructs might be represented by labels such as kind–cruel, shy–outgoing or happy–sad. Through a system of these personal construct dimensions an individual is able to predict and control his interpersonal world. Research has focused on the investigation of the vicissitudes of personal construct systems as determinants of behaviour (Bannister & Mair, 1968; Bonarius, 1965). The immediate task of much of the research has been to obtain an empirical assessment of the constructs in a form consistent with theoretical specifications.

Kelly (1955), however, described two different formal methods for eliciting constructs in bipolar form. The first method, labelled the difference method (DM), was presented in the individual interview format of the repertory grid technique. The second method, referred to here as the opposite method (OM), was presented in the group administration format. In both methods the likeness end of the construct is elicited in the same way. It is the word or short phrase given to describe the similarity when a subject is asked to think of how two of three people are alike in some important way and, at the same time, different from the third. The two methods differ in the way in which the contrast end is elicited. In the DM the subject is asked to give the word or short phrase that describes how the third person is different from the other two. Using the OM the subject is asked to give what he considers to be the opposite of the characteristic he listed for the likeness end. General descriptions of the theory (Bannister & Mair, 1968) and some research studies have used the difference method (Cromwell & Caldwell, 1962; Flynn, 1959). Other investigators have used the opposite method (Levy, 1956; Levy & Dugan, 1956; Fager, 1958) and some have used a combination of the two methods (Abma, 1969; Landfield, 1954).

Since bipolarity is of central importance in the definition of a construct and since it can affect the interpretation of results obtained in studies of construct systems, the question is raised here concerning what possible difference might exist between the two methods as operations for obtaining bipolar constructs. While other investi-

gations have been carried out on the oppositional or bipolar nature of constructs (Resnick & Landfield, 1961; Mair, 1967), there has not been a direct investigation comparing the bipolarity of specific methods.

Method

The subjects were 46 female undergraduates who were fulfilling a psychology course requirement at the University of Florida. Owing to the variation in the size of the groups in which subjects signed up for the experiment, 22 subjects received the OM instructions and 24 subjects received the DM instructions. Three subjects had to be discarded for failing to follow instructions, leaving 20 subjects in the OM group and 23 subjects in the DM group.

First session: instruments and procedure

During the first session, subjects completed a modified (22 × 12) role construct repertory grid. The standard role list of 22 personal acquaintances (referred to as elements) appeared across the top of the grid. Each of the 12 rows of the grid designated three elements (referred to as a sort) from the 22 person role list which were used to elicit the 12 constructs. The triad of elements composing each of the 12 sorts was chosen from those suggested by Kelly (1955).

All subjects were given identical instructions for eliciting the likeness end of the constructs. The subjects were told to 'Think about the three people. Are two of them alike in some important way that distinguishes them from the third person? Write the word or short phrase that tells how these two are alike.' The two experimental conditions differed in the way the contrast end was elicited. The subjects in the OM condition were asked to give the 'opposite' of the word or short phrase initially elicited, while the subjects in the DM condition were asked to give the 'way the third person is different'.

Second session: instruments and procedure

In session two, all subjects were individually administered modified grids with the 22 elements used in the first grid appearing across the top. The purpose of this second grid was to assess the bipolarity of the constructs elicited in the first grid. The mean time between sessions was 5 days. In this second grid, no sorts were indicated for generating constructs. The constructs were all single-pole construct labels provided by the experimenter. The provided constructs were the single poles of the subject's 12 constructs elicited in the first session. The single-pole labels were placed at the right of each row, requiring 24 rows for each grid. The likeness end labels and the contrast end labels were grouped separately. Within their respective groupings, the order was random and the two groupings were counterbalanced for the total subject pool. For each of the 24 single-label constructs, the subject was instructed to consider the 22 elements and place a check mark in exactly one-half (11) of the cells in a row for elements to which that label applied. This technique of grid administration, developed by Bannister (1960), is termed the split-half form of the repertory grid.

After a subject completed the grid in session two, interview questions were asked in order to assess the subject's ability to verbalize the contrasting quality of bipolar constructs. Using a modification of Hunt's (1951) procedure, the questions were designed to set a construct's two poles in conceptual opposition in order to ascertain if they were the contrasting poles of a single construct. The experimenter read the following question to the subject and recorded the subject's 'yes' or 'no' response. 'Is a person who is (likeness end of a construct) almost never (contrast end of the contruct)?' The question was repeated with the reverse order of labels. These questions were asked for each of the 12 elicited construct dimensions providing 24 responses; a 'yes' answer indicated contrasting poles of a single bipolar construct.

Scoring for bipolarity in grids

Bipolarity is assessed by the number of matches between the two rows in the second grid containing the end labels of the original construct dimension. The scoring procedure consists of taking the two rows of interest and counting the number of exact matches. Taken in steps of two, the total range of matches is from 0 to 22. For rows 22 cells in length, the significance of

a matching score can be calculated by the method of binomial expansion. With $P < 0{\cdot}05$, a match of 16 or more indicates a positive relationship or overlap between the two end labels. A match of six or less indicates a negative relationship between the two end labels, which is the inverse relationship expected between the poles of the true bipolar constructs. Matches between 8 and 14 represent a chance relationship.

Results

A difference between the two methods is demonstrated by examining the proportion of constructs showing the bipolar negative relationship. Of the 240 constructs obtained from the 20 subjects in the OM condition, 175 (72·9 per cent) match six or fewer times. Of the 276 constructs elicited from the 23 subjects receiving DM instruction, 156 (56·5 per cent) met the criterion for bipolarity. The difference between these two proportions is significant ($z = 3{\cdot}88$; $P < 0{\cdot}001$). An examination of the remaining constructs revealed that the poles of 10 constructs (4·2 per cent) from the OM were overlapping, while 55 (22·9 per cent) showed no relationship. Fifteen (5·5 per cent) DM constructs overlapped and 105 (38·0 per cent) indicated no relationship between poles.

A more demanding test of bipolarity would be to require an almost perfect reflexion between elicited construct labels. Consequently, the number of constructs with zero or two matches between respective poles was noted. Seventy-nine (33 per cent) of the constructs from the OM and 47 (17 per cent) of the constructs from the DM fell into this category, yielding a significant difference between the proportions ($z = 4{\cdot}22$; $P < 0{\cdot}001$).

Without regard to the significance of the matches between the two pole labels, another index of bipolarity can be obtained by counting the total number of matches occurring between the two rows. A small number of matches indicates a high degree of bipolarity. In comparing the two methods, the subjects receiving the OM instructions produced a lower mean of 64·6 matches when compared to the DM mean of 83·7 ($t = 2{\cdot}98$; $P < 0{\cdot}005$).

The data obtained from the second session interview questions were incomplete. Only 12 subjects in the OM group and eight subjects in the DM group responded with 'yes' or 'no' to all of the 24 questions asked. Some of the constructs of the remaining subjects appeared meaningless when placed in the standard form and subjects were unwilling to give clear 'yes' or 'no' responses. Since 240 constructs were elicited from the OM group, the standard question was posed 480 times. Likewise, there were 522 potential 'yes–no' responses from the DM group. The OM subjects could answer 418 (87·1 per cent) of these questions while the DM subjects answered only 406 (73·7 per cent). Based only on responses obtained, the difference in the distributions of 'yes' *v.* 'no' responses from the two methods failed to reach significance ($\chi^2 = 3{\cdot}00$; $P < 0{\cdot}10$).

Discussion

Because responses to the interview technique were incomplete, this discussion will deal specifically with the results from the grid technique. Using the formal methods for eliciting bipolar constructs to be employed in discrimination tasks involving interpersonal events, the OM appears to have an advantage over the DM. These results are limited, however, to the specific characteristic of bipolarity. The relative

advantage or disadvantage of the two methods might be quite different if other characteristics of constructs were of primary interest. In addition, the yield of bipolar constructs from the more efficient OM indicates that improvement, even in this method, is needed in order to obtain a total sample of clearly bipolar constructs for each subject. Although 72·9 per cent of the constructs obtained using the OM reached the $P < 0 \cdot 05$ criterion of six or less matches, only 33 per cent reached the more demanding criterion of zero or two matches.

A number of explanations might be offered in order to account for the results obtained in this study. For example, it is possible that attributes of the DM instructions produced the difference between the two methods. These instructions perhaps caused the subject to shift to a separate construct dimension in order to formulate a contrast pole. Instead of giving the contrast pole, the subject could be naming a separate attribute of the third element; perhaps the likeness end label of another construct. In addition, the DM instructions could have the effect of producing constructs with restricted ranges of convenience. Because the contrast pole must specifically represent the third element, the construct formed would be tailored to these elements and not applicable to other elements. The occurrence of either or both of these mediating processes would result in overlap between construct poles.

The focus of the present research is on the demonstration that differences in bipolarity occur when the two methods are employed in an element sorting task. Further research is needed both to offer a way of differentiating among possible explanations for the present results and to provide improvements in the existing formal methods for eliciting bipolar constructs.

The present experiment was based in part on a master's thesis by C. J. Nickeson and was presented at the annual meeting of the Southeastern Psychological Association, Louisville, Kentucky, April 1970. Requests for reprints should be sent to Franz R. Epting, Department of Psychology, University of Florida, Gainesville, Florida 32601.

References

Abma, B. (1969). Changes in construction of parent and child figures following enactment: playing a role. (Doctoral dissertation, Ohio State University. Ann Arbor, Michigan: University Microfilms.)

Bannister, D. (1960). Conceptual structure in thought-disordered schizophrenics. *J. ment. Sci.* **106**, 1230–1249.

Bannister, D. & Mair, J. M. M. (1968). *The Evaluation of Personal Constructs.* New York: Academic Press.

Bonarius, J. C. J. (1965). Research in the personal construct theory of George A. Kelly: role construct repertory test and basic theory. In B. A. Maher (ed.), *Progress in Experimental Personality Research*, vol. 2. New York: Academic Press.

Cromwell, R. L. & Caldwell, D. F. (1962). A comparison of ratings based on personal constructs of self and others. *J. clin. Psychol.* **18**, 43–46.

Fager, R. E. (1958). Student and faculty conceptions of 'successful student'. *J. counsel. Psychol.* **5**, 98–103.

Flynn, J. C. (1959). Cognitive complexity and construct constellatoriness as antecedent conditions of role variability. (Unpublished master's thesis, Ohio State University.)

Hunt, D. E. (1951). Studies in role concept repertory: conceptual consistency. (Unpublished master's thesis, Ohio State University.)

Kelly, G. A. (1955). *The Psychology of Personal Constructs*, vols. 1–2. New York: Norton.

Landfield, A. W. (1954). A movement interpretation of threat. *J. abnorm. soc. Psychol.* **49**, 529–532.

Levy, L. H. (1956). Personal constructs and predictive behaviour. *J. abnorm. soc. Psychol.* **53**, 54–58.

Levy, L. H. & Dugan, R. D. (1956). A factorial study of personal constructs. *J. consult. Psychol.* **20**, 53–57.

Mair, J. M. M. (1967). Some problems in repertory grid measurement. I. The use of bipolar constructs. *Br. J. Psychol.* **58**, 261–270.

Resnick, J. & Landfield, A. W. (1961). The oppositional nature of dichotomous constructs. *Psychol. Rec.* **11**, 47–55.

(*Manuscript received* 10 *November* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 519–525
Printed in Great Britain

WORD ASSOCIATION, ASSOCIATIVE STRUCTURE AND MANIFEST ANXIETY

By JOHN M. INNES

Department of Psychology, University of Birmingham

Undergraduate subjects gave continuous associations to four stimulus words varying in connotation and abstractness. Subjects scoring high and low on the Manifest Anxiety Scale were selected; these subjects then gave continuous associations to their first five responses to each of the original stimulus words. A measure of inter-item associative strength was obtained for each of the original stimuli. The highly anxious subjects' associations to the words with positive connotation were found to be more cohesive than these subjects' associations to the words with negative connotation. No such difference was found for low anxiety subjects nor was any difference found between associative structures for words varying in abstractness. The results were taken to suggest that the structure of associations to stimulus words may be an important variable in accounting for personality differences in word association behaviour.

In recent years several studies have been conducted with the object of demonstrating the role of inter-word associations among a particular set of words, i.e. the structure of associations, in the generation of associations in a discrete or in a continuous word-association task. Much of this work is associated with Pollio (1968, 1966, 1964). Pollio has attempted to show that the immediate and higher-order linkages which can be calculated for a restricted set of words by means of matrix algebra (Pollio, 1963) relates to the speed of emission of word associations in continuous association.

What Pollio has shown in effect is that words can be inferred to exist in structures and the 'distance' between these words can be estimated in terms of the number of words that intervene associatively between two particular words. Using sets of words in the form of a matrix where each row is headed by one of the words and each column is also headed by a word, one can develop $n \times n$ matrices, with the cells of the matrix containing either 1 or 0, indicating that a column word is elicited by a row word, or not. Raising these matrices through the first to the nth power one can calculate the number of first, second, third, ..., nth order steps that connect two words. Summing the multiplied matrices algebraically one can arrive at an estimate of the total associative 'connectedness' between the words. Pollio (1964) gives a detailed description of the method.

Pollio (1964) has shown that words which in continuous association are emitted with short inter-word latencies have more restricted inter-item structure than words which are emitted with long between-word latencies.

The present experiment was undertaken to see if the structures produced by individual subjects could be related to personality differences previously shown to relate to word-association behaviour. The variable studied was anxiety, as measured by the Taylor Manifest Anxiety Scale (MAS), taken as a measure of drive.

Hull–Spence theory predicts that an increase in drive should increase the probability of emission of a response which is high in the habit family hierarchy of responses to a stimulus. If the MAS is a measure of the 'arousability' of an emotional response

to a test situation, then subjects scoring high on the MAS should have higher probabilities of emitting primary responses and their responses should have shorter latencies than low-scoring subjects.

In word association high scoring MAS subjects should give faster responses, and assuming that cultural frequencies as listed in the norms reflect individual hierarchy strength, they should give more primary responses than low-scoring subjects. In a discrete free-association task Kanfer (1960) found no evidence to support this hypothesis. High and low MAS scorers did not differ in the frequency of primaries to words from the Kent–Rosanoff list. Some of the literature indicates that moderately anxious subjects give more primaries (Sarason, 1959, Wolff, 1965).

High scorers do produce more different responses than do low scorers. Davids & Eriksen (1955) in a chain-association task found a significant positive correlation between MAS scores and number of associations. Johnson & Lim (1964) found high scorers to produce more continuous associations to words rated as 'good' than to 'bad' stimuli than did low scorers. This latter study may also point to a factor which may account for any contradictions in the literature. The emotional value of a stimulus word may have to be controlled before any clear difference emerges.

Invoking the idea of associative clusters of interconnected words, we may predict that high MAS scorers should have more tightly clustered structures of associations than low scorers. Each response word acts as a stimulus for other implicit associates and these links should be enhanced by a high drive state. The effect of increasing drive, or the subject having a greater predilection to drive arousal, should be to produce highly interconnected structures, and this should be reflected in faster generation of associations in overt continuous association.

The present experiment is concerned with demonstrating that high MAS scorers have more dense word-association structures underlying the associative behaviour and produce more associations in continuous word association. This effect should be especially observable among words known to have tight associative structures, thus high MAS subjects should have denser structures than low scorers for words normally highly interconnected.

From previous work by Pollio (1964, 1966) it appears likely that words with positive as against negative connotation are more densely clustered. Previous work has also shown that words judged as 'concrete' have more associates than words judged as 'abstract' in meaning (Pollio, 1964, 1966; Lambert, 1955; Paivio *et al.*, 1968). To the degree that a larger number of associations indicates a denser associative domain then the prediction might be made that high MAS scorers should have more densely related associative structures for positive compared to negative words than low MAS scorers, and also for concrete compared to abstract words.

METHOD

The experiment involved two sessions for the subjects. In the first session the subjects associated continuously to four different stimulus words for 2 min. each and also completed a 20-item form of the Manifest Anxiety Scale (Bendig, 1956). After this subjects scoring high or low on the MAS were selected and presented with the five words they had produced at the beginning of each association period for all four stimulus words. Both sessions were conducted in group conditions. The sessions were 2 weeks apart.

Subjects. For session 1 a total of 36 male and 46 female undergraduates were tested. From the MAS data of session 1, 12 male and 12 female subjects were selected who had high and low scores on the anxiety scale. Six men and six women were high scorers and an equal number of men and women were low scorers. Subjects were selected in rank order of score from the top and from the bottom of the distribution of scores, the scores for men and women being ranked separately.

The means and standard deviations for the total samples were $5 \cdot 8 \pm 3 \cdot 9$ for men and $5 \cdot 9 \pm 3 \cdot 1$ for women. These data conform closely to data reported by Howe (1960). For the extreme groups selected for the continued word-association task, the means for the high and low groups were 11·1 and 2·0 respectively for men and 11·5 and 2·5 for women.

Stimuli. Four words were chosen as stimuli: DOCTOR, JUSTICE, THIEF and TROUBLE. These words were taken from Pollio's (1964) set of words except for DOCTOR which was substituted for HOUSE. The words were chosen on the following criteria: (*a*) their evaluative semantic-differential rating and (*b*) their frequency of occurrence in English. Table 1 gives the four words and their characteristics. Two words were considered to be concrete (DOCTOR and THIEF) and two to be abstract (JUSTICE and TROUBLE).

Table 1. *Stimulus words and their characteristics*

Stimulus word	Evaluative rating*	Rated goodness†	Frequency of occurrence‡	Concreteness§	Concreteness†
Doctor	1·87	6·21	AA	6·62	6·47
Justice	2·20	6·31	A	2·18	3·03
Thief	5·59	1·60	28	6·08	6·00
Trouble	5·03	2·00	AA	2·25	3·89

* From Jenkins (1960).
† From Brown & Ure (1969).
‡ From Thorndike–Lorge (1944) G count.
§ From Paivio *et al.* (1968).

Data from Brown & Ure (1969) and from Paivio *et al.* (1968) which became available some time after the completion of the experiment are presented as confirmation of the earlier categorization of abstractness.

Materials and procedure. Four-page booklets were made up with a stimulus word to each page. The word was printed at the top of the page and repeated down the page in two columns. Booklets were made up in all 24 possible orders and were distributed to subjects in batches so that each set of orders was completed before the next set was begun.

The instructions were as used in Noble's (1952) study. Subjects were told they would have 2 min. for each word. Each page was timed separately.

After the association test each subject completed a 20-item form of the MAS.

A separate word-association task was constructed for each subject. The first five response words to each of the four stimulus words were selected. Each word was printed at the top of a separate sheet of paper and then repeated five times down the side. A blank space appeared to the right of each stimulus word where the subject had to place his response. The order of words in the booklet was randomly constructed, with the restriction that not more than two words from the same sequence should follow each other in any booklet. For each subject therefore there was a specific booklet constructed from his own associations. Instructions were as for the first session, although subjects this time proceeded through the booklet at their own speed.

Method for measuring associative cohesiveness. For each subject, for each of the four initial stimulus words, a measure of total potential cohesiveness (MC) was computed giving the degree of cohesiveness that existed for that subject for that stimulus word. The method for obtaining MC values is described in detail in Pollio (1964). In Pollio's study a measure of relative cohesiveness was constructed as his matrices were of different sizes. In the present study, however, all matrices were 5×5 so it was not considered necessary to compute relative values. In order to compare the present results with those of Pollio (1964), it is only necessary to divide the MC values reported here by 25.

RESULTS

Associative cohesiveness (MC). Table 2 gives the means and standard deviations for the MC values as a function of connotation and anxiety. An analysis of variance was performed on the MC data after a $\sqrt{(x + 0{\cdot}5)}$ transformation, as the data showed marked skewness with several values being zero. The analysis took the form of a $2 \times 2 \times 2 \times 2$ design, with anxiety and sex as between subjects factors and connotation and abstractness as within subjects factors.

The analysis indicated connotation to be a significant factor ($F = 6{\cdot}97$, d.f. = 1, 20, $P < 0{\cdot}05$). Positively evaluated words produced associations which were more cohesive than words which were negatively evaluated.

Table 2. *Means and* S.D.*s of index of cohesiveness* (MC) *as a function of connotation and anxiety*

	Connotation			
	Positive		Negative	
Anxiety	Mean	S.D.	Mean	S.D.
High	4·49	1·87	2·87	0·94
Low	3·96	1·69	3·85	1·30

Table 3. *Means and* S.D.*s of number of associations for connotation and anxiety conditions*

	Connotation			
	Positive		Negative	
Anxiety	Mean	S.D.	Mean	S.D.
High	14·00	4·75	11·71	3·84
Low	11·58	4·61	11·08	4·46

The connotation × anxiety interaction was also significant ($F = 5{\cdot}44$; d.f. = 1, 20; $P < 0{\cdot}05$). High-anxiety subjects had more cohesive associative clusters with positively evaluated words than with negative words, low-anxiety subjects showing no difference. Individual comparisons of the means at high anxiety (Winer, 1962, p. 611) showed that the difference between positive and negative connotations was significant. The differences within low anxiety were not significant. Comparisons of means between anxiety levels showed no significant differences.

No other effect approached a conventional level of significance. Concrete words were no more structured than abstract and high- and low-anxiety subjects did not differ in the degree of structuring.

Continued association (M). The total number of associations given to each stimulus word by the subjects was found and an analysis of variance of the same form as that performed on the MC data was carried out.

Positively evaluated words elicited more associations than negatively evaluated words ($F = 9{\cdot}17$; d.f. = 1, 20; $P < 0{\cdot}01$). Concrete words also elicited more associations than abstract ($F = 14{\cdot}98$; d.f. = 1, 20; $P < 0{\cdot}01$). The means were 13·31 and 10·88 for concrete and abstract words respectively.

The anxiety × connotation interaction approached a conventional level of significance ($F = 3{\cdot}78$, d.f. $= 1, 20$, $0{\cdot}10 > P > 0{\cdot}05$). The means for this interaction are given in Table 3.

High-anxiety subjects apparently produced more associations to positively evaluated than to negatively evaluated words and also more compared to low-anxiety subjects. Such an effect did not appear for concrete as opposed to abstract words.

Discussion

The data give some support for the hypothesis that positively evaluated words have more cohesive structures, as indexed by MC values, than negatively evaluated words. The other finding, that concrete words do not have more cohesive structures than abstract words, however, does not support the experimental hypothesis.

High-anxiety subjects apparently have more cohesive structures relating positively evaluated words than they have for negatively evaluated stimuli. Low-anxiety subjects show no difference between these classes of stimulus terms. It seems possible that high-anxiety subjects have 'tighter' associative structures for words which, from normative data already available, have more dense associative relationships.

The data also indicate that subjects give more associations, as indexed by M, to positively evaluated stimuli. This finding supports previous work by Bousfield (1944), Johnson & Lim (1964), Koen (1962) and Pollio (1964, 1966). It also seems that high-anxiety subjects give more associations to positive than to negative evaluative words. It may be, therefore, that the production of associations in a continued word-association task is dependent upon the cohesiveness of the relationships between the associations, and that the effect of anxiety is to increase cohesiveness and thus associative production.

The data on the concrete–abstract dimension, however, does not support such an interpretation entirely. While subjects gave more associations to concrete as compared to abstract nouns, thus supporting previous work by Lambert (1955) and Paivio *et al.* (1968), concrete words did not have more cohesive structures than abstract words, nor was there an interaction of Anxiety with Abstractness on either the MC or the M values. Such failures do not support the idea that meaningfulness is a function of the stimulus words mediated by the cohesiveness of the structure relating associations of the stimulus word.

Paivio *et al.* (1968) have suggested that concrete and abstract nouns differ in the nature of the associations available to them. Abstract nouns may have meaning (as indexed by M) because of the relationships between verbal associations. Concrete nouns in addition may have sensory and imaginal associations. Sensory associations may break up the verbal associative structures of concrete nouns, producing relatively diffuse verbal associative networks. Abstract nouns may have smaller associative domains with structures which are predominantly verbal but which are not absolutely more cohesive than those of concrete nouns because of the relative paucity of associations. Such a different basis for the production of associations to concrete and abstract nouns may explain the failure to find differences in cohesiveness related to differences in meaningfulness. Abstract nouns did tend to have higher MC values than concrete, although the difference was far from being significant.

It may be possible, of course, that much more than the 'normative' associative structure of connotative material is affected by the presence of anxiety as a subject variable. High-anxiety subjects may respond to negatively evaluated material in a qualitatively different manner from low-anxiety subjects. Negatively connotative words may elicit neutral or anxiety-reducing responses which would perhaps result in associative structures which are not highly related. Examination of the data suggests that high-anxiety subjects do not show only a tightening of structure for positively evaluated nouns, but also a loosening of structure for negatively evaluated nouns. High-anxiety subjects do seem to have lower *MC* values for negative stimuli than do low-anxiety subjects, although this difference is not statistically significant.

The present study does suggest, however, that differences in the associative structures underlying associative production may exist between subjects varying in anxiety. Further work with a wider range of stimulus materials, varying *MC* values over a greater range, would be valuable. Also important would be the computation of *MC* values for associations from later on in the production time. After 5–10 min. of association, when subjects are exploring more widely into their associative domains, measures of the structure of these peripheral domains might indicate differences in the use of mechanisms by high- and low-anxiety subjects dealing with positive and negative evaluative material.

References

BENDIG, A. W. (1956). The development of a short form of the Manifest Anxiety Scale. *J. consult. Psychol.* **20**, 384.

BOUSFIELD, W. A. (1944). An empirical study of the production of affectively toned items. *J. gen. Psychol.* **30**, 205–215.

BROWN, W. P. & URE, D. M. J. (1969). Five rated characteristics of 650 word association stimuli. *Br. J. Psychol.* **60**, 233–249.

DAVIDS, A. H. & ERIKSEN, C. W. (1955). The relationship of manifest anxiety to association productivity and intellectual attainment. *J. consult. Psychol.* **19**, 219–222.

HOWE, E. S. (1960). Further comparison of two short-form derivatives of the Taylor Manifest Anxiety Scale. *Psychol. Rep.* **6**, 21–22.

JENKINS, J. J. (1960). Degree of polarization and scores on the principal factors for concepts in the semantic atlas study. *Am. J. Psychol.* **73**, 274–279.

JOHNSON, R. C. & LIM, D. (1964). Personality variables in associative production. *J. gen. Psychol.* **71**, 349–350.

KANFER, F. H. (1960). Word association and the drive hypothesis of anxiety. *J. clin. Psychol.* **16**, 200–204.

KOEN, F. (1962). Polarization and emotionality in words. *J. verb. Learn. verb. Behav.* **1**, 183–187.

LAMBERT, W. E. (1955). Associational fluency as a function of stimulus abstractness. *Can. J. Psychol.* **9**, 103–106.

NOBLE, C. E. (1952). An analysis of meaning. *Psychol. Rev.* **59**, 421–430.

PAIVIO, A., YUILLE, J. C. & MADIGAN, S. (1968). Concreteness, imagery and meaningfulness values for 925 nouns. *J. exp. Psychol. Monogr. Suppl.* **76**, no. 1, part 2.

POLLIO, H. R. (1963). A simple matrix analysis of associative structure. *J. verb. Learn. verb. Behav.* **2**, 166–169.

POLLIO, H. R. (1964). Composition of associative clusters. *J. exp. Psychol.* **67**, 199–208.

POLLIO, H. R. (1966). *The Structural Basis of Word Association Behaviour*. Hague: Mouton.

POLLIO, H. R. (1968). Associative structure and verbal behaviour. In T. R. Dixon & D. L. Horton (eds.), *Verbal Behaviour and General Behaviour Theory*. Englewood Cliffs: Prentice-Hall.

SARASON, I. G. (1959). Relationships of measures of anxiety and experimental instructions to word association performance. *J. abnorm. soc. Psychol.* **59**, 37–42.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Columbia University Press.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

WOLFF, C. (1965). Manifest anxiety, reaction potential ceiling and word association. *J. Person. soc. Psychol.* **2**, 570–573.

(*Manuscript received* 20 *August* 1970; *revised manuscript received* 8 *October* 1970)

Br. J. Psychol. (1971), 62, 4, *pp.* 527–531
Printed in Great Britain

STIMULUS COMPLEXITY AND THE OCCIPITAL EEG

BY ANTHONY GALE, BRUCE CHRISTIE AND VALERIE PENFOLD
Department of Psychology, University of Exeter

Occipital EEG was monitored during long-term exposure to visual stimulus arrays of five levels of complexity. Integrated EEG output was recorded for nine separate frequency bands. The stimuli were two, four, eight, 16 or 32 randomly located white squares on a black background. Mid-alpha activity (8·5–10·5 Hz) decreased linearly with $\log_2 n$, where n = number of squares in the array ($P < 0{\cdot}01$). Beta activity (12·5–16·5 Hz) showed a quadratic trend ($P < 0{\cdot}02$). Theta activity (5·5–7·5 Hz) increased linearly with ascending complexity ($P < 0{\cdot}05$). There were no significant trends for low alpha (7·5–8·5 Hz) or high alpha (10·5–12·5 Hz). Brightness-control slides showed no parallel effects.

A number of studies have demonstrated a relationship between the EEG and complexity of stimulation (Berlyne & McDonnell, 1965; Baker & Franken, 1967; Gale *et al.*, 1969*b*). They support the view that lower amplitude EEG is associated with higher complexity. The present study (i) extends the range of EEG frequencies; (ii) employs 'number of display elements' as a measure of complexity; (iii) includes brightness-control stimuli; and (iv) samples EEG automatically over relatively long periods of visual stimulation. This study therefore combines the better features of earlier work and provides a basis for a more parametric approach.

METHOD

Subjects. The subjects were eight male and six female students at the University of Exeter, with an average age of 18 years 6 months.

Apparatus, procedure and statistical analysis

Displays. A sample of experimental and control displays is shown in Fig. 1. The experimental displays contained two, four, eight, 16 or 32 white squares distributed at random on an 8 × 8 black matrix. Adjacent squares were separated by contour to reduce the formation of other shapes by fusion. To reduce the possible effects of particular configurations, six such sets were constructed, all differing in the random location of the square elements. The control displays consisted of single central squares of total areas equivalent to two, four, six, eight, 16 and 32 elements. Thus control was provided for brightness, area and contour-length. The displays were back-projected on to a 22 in. square screen (tracing paper sandwiched between glass) to give a 13 in. square display in which the square elements were approximately 1·45 in. and the contour line between adjacent squares was 0·2 in. Illumination at the subject's distance from the screen (7·5 ft.) varied for both experimental and control slides between 1 and 2 LM ft.^{2}, in equal steps depending on total area of white. The remainder of the screen and the experimental room were in total darkness.

The EEG. Silver/silver chloride loose-pad electrodes were placed transoccipitally for bipolar recording (Cooper *et al.*, 1969, p. 78). Mean inter-electrode resistance was 6·8 kΩ and never exceeded 18 kΩ. The neutral electrode was placed on the right wrist. The EEG was recorded on a San'ei polygraph calibrated to give 28 mm write-out (peak to trough) for 100 mV. The time constant, set at 0·1 sec., was optimal for the range measured. The EEG was integrated by a San'ei Low Frequency Analyser over nine separate pure flat-top band-pass filters, with steep skirts and linear output, employing an epoch time of 5 sec. The filters ran from 5·5 to 13·5 Hz in single frequency steps (eight filters) with a ninth filter at 13·5–16·5 Hz. Prior to each experimental run, each filter was tested with an oscillator set at the midpoint for its frequency. Both

primary recording and integrated output were written out on continuous millimetre recording paper run at 5 mm per sec. At the same time the integrated outputs of the filters were punched out in a six-hole binary code (maximum resolution of 63 units) by a Lion Systems Development A-D converter. Onset of each trial was also coded manually on the punch-tape as was any gross movement artifact on the primary record. The binary code information was then decoded to decimal numbers by the university computer.

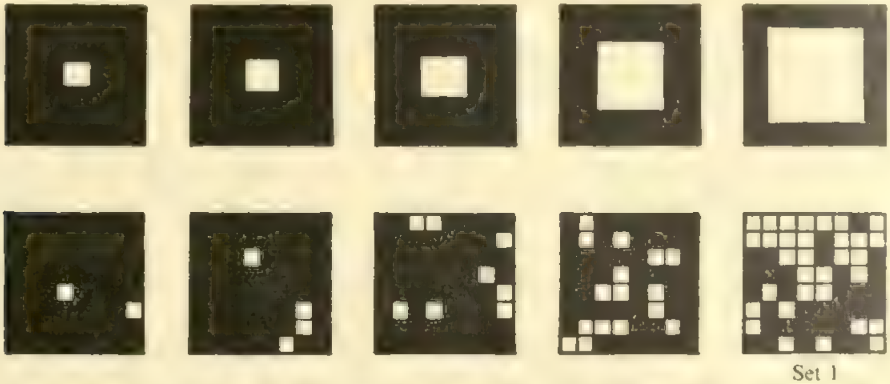


Fig. 1. A complete set of control and experimental slides.

General design. The subjects were divided into two groups of four males and three females. Group 1 saw experimental display sets 1–3 (five complexity levels per set), with three sets of control displays. Thus there were 30 slides per subject: 15 experimental, 15 control. Group 2 saw sets 4–6 plus controls. Both the order of experimental sets and the order of displays within sets were randomized for each subject. Each experimental slide was presented for 70 sec. and preceded or succeeded (on the basis of a random schedule) by its associated control slide (also for 70 sec.). Order of presentation was coded until data reduction was complete.

Procedure. The subject sat on a hairdresser's chair the height of which was adjusted until his line of regard was at 90° to a central fixation point on a test slide. The EEG electrodes were applied and a sample recording taken. This setting-up procedure (which was carried out under the dull illumination of the test slides) took on average 12 min. The subject was instructed to observe the slides passively without counting the squares or attempting to estimate the passage of time. He was informed that no questions would be asked concerning the slides. He was to attend to the displays, to keep his eyes open and to attempt to avoid distraction. It was not possible to mask entirely the sound of the paper punch in the adjacent room but its sound was constant for all conditions. The slides were then presented according to the particular subject's schedule. On the few occasions when gross artifacts occurred, the trial was extended by an extra 5 sec. epoch. The experimenter could monitor the display screen on each trial by means of a masked spy-hole. Total running time was approximately 35 min.

Data reduction. For each subject the line-print contained integrated output data for 30 slides (14 epochs per slide, nine filter values per epoch). Since stimulus presentation did not coincide with epoch-onset the 10 central epochs were selected (excluding the first and final pairs). The data for the three display sets were collapsed for each of the nine filters, for each of the five complexity levels (both experimental and control), yielding 90 scores per subject.

Statistical analysis. The technique outlined by Still (1967) was employed. The output from each filter was tested for linear, quadratic and cubic trends, experimental and control displays taken separately. A trend test was appropriate since the interval between experimental conditions (i.e. number of display elements) was known. To avoid calculation of orthogonal polynomial coefficients for unequal intervals (Gaito, 1965), the independent variable was expressed as $\log_2 n$ (where n = number of squares in the experimental displays or area of central square in the control displays). For each subject and each filter the five experimental and control scores were multiplied by the corresponding orthogonal polynomial coefficients for equal intervals (Fisher &

Yates, 1963), and the resulting five weighted values were summed. This was done for each of the trend components: linear, quadratic and cubic.

A comparison was made between groups 1 and 2 to ensure that any complexity effect would not be attributable to a *specific* set of slides. The trends for all conditions were compared using the Mann–Whitney U test (Siegel, 1956). Of 54 such tests only three attained significance at the 5 per cent level.

The significance of trends was tested with the Wilcoxon T (Siegel, 1956), collapsing groups 1 and 2, to determine whether the weighted sums were greater or less than zero. Where groups 1 and 2 had been shown to differ, test for trends was also performed separately.

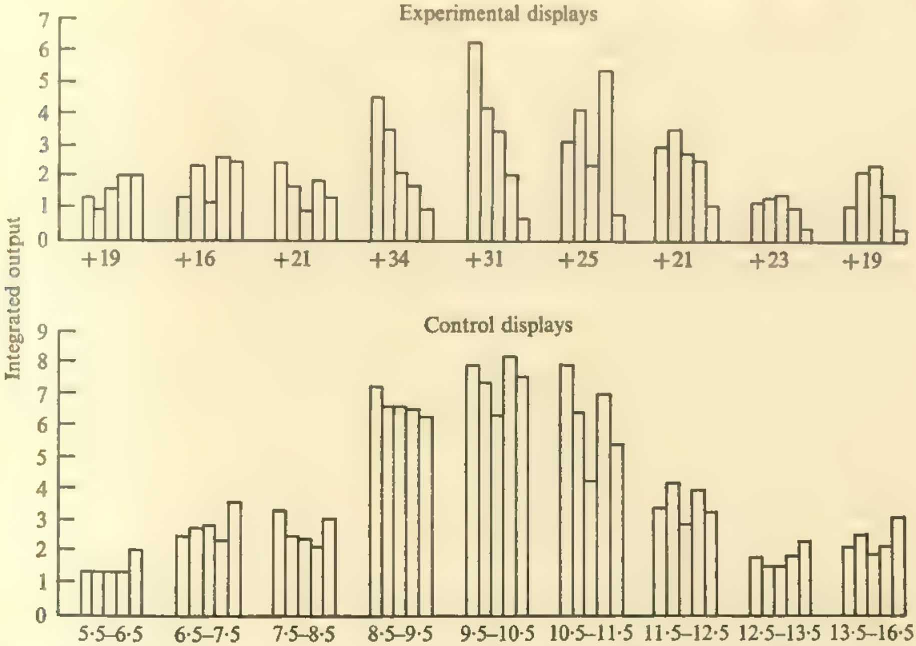


Fig. 2. Integrated output for each of the five complexity and brightness levels (left to right in ascending complexity). Since output varies considerably for the different frequency bands, different base lines are used for different filters. For true value, add the number at the base of the graph (same for control and experimental slide) to the value on the ordinate. All the graphs are drawn to the same scale and statistical tests performed on filters separately.

Table 1

	Filter								
Trend	5·5–6·5	6·5–7·5	7·5–8·5	8·5–9·5	9·5–10·5	10·5–11·5	11·5–12·5	12·5–13·5	13·5–16·5
				Experimental displays					
Linear	$P < 0{\cdot}05$	$P < 0{\cdot}05$	—	$P < 0{\cdot}01$	$P < 0{\cdot}01$	—	—	—	—
Quadratic	—	—	—	—	—	—	—	$P < 0{\cdot}01$	$P < 0{\cdot}02$
Cubic	—	—	—	—	—	—	—	—	—
				Control displays					
Linear	—	—	—	—	—	—	—	—	—
Quadratic	—	—	—	—	—	$P < 0{\cdot}02$*	—	—	—
Cubic	—	—	—	—	—	—	—	—	$P < 0{\cdot}02$

* Group 1 only.

Results

A summary of the trend data is shown in Table 1. Mean integrated outputs for the 14 subjects for all filters and experimental and control conditions are shown in Fig. 2.

Discussion

Previous studies of the relation between the EEG and stimulus complexity have been largely concerned with short-term response to brief stimulus exposure (Berlyne & McDonnell, 1965; Baker & Franken, 1967). The present study demonstrates discrete variation in EEG during relatively long-term exposure to the stimulus. Previous work has focused on alpha frequencies. We confirm previous findings that mid-alpha abundance decreases with increased complexity. The reverse finding for theta frequencies was not expected, and calls for more systematic study. Daniel (1967) also reports higher amplitude theta at times when alpha attenuation might be expected (i.e. prior to detections in a vigilance task when compared with theta prior to omissions). We have discussed elsewhere some of the difficulties of interpreting changes in theta activity in human subjects (Gale *et al.*, 1969*a*, pp. 220–221). Again, the quadratic trend for beta, which holds for two distinct frequency bands, calls for a replication study. We propose to extend the range and number of stimuli to see if the effect holds for a much greater number of stimulus categories, i.e. whether maximum activity always appears at a central point. In the absence of other studies of this frequency range the robustness of the finding must be suspect.

As in the study of Baker & Franken (1967), brightness, and presumably size and contour length, have relatively little effect.

The notion of 'complexity' is of course a dubious one in the present context. Berlyne (1960) considers the problem of complexity and its definition in some detail. However, on an intuitive level, number of elements in the display would appear to be a good candidate for *one* aspect of complexity at least. Certainly, we have demonstrated systematic occipital EEG change with systematic manipulation of the stimulus.

This work was supported by a grant from the Medical Research Council (reference G969/185/C).

References

Baker, G. & Franken, R. (1967). Effects of stimulus size, brightness and complexity upon EEG desynchronization. *Psychon. Sci.* **7**, 289–290.

Berlyne, D. E. (1960). *Conflict, Arousal and Curiosity*. New York: McGraw-Hill.

Berlyne, D. E. & McDonnell, P. (1965). Effects of stimulus complexity and incongruity on duration of EEG desynchronization. *Electroenceph. clin. Neurophysiol.* **18**, 156–161.

Cooper, R., Osselton, J. W. & Shaw, J. C. (1969). *EEG Technology*. London: Butterworth.

Daniel, R. S. (1967). Alpha and theta EEG in vigilance. *Percept. mot. Skills* **25**, 677–703.

Fisher, R. A. & Yates, F. (1963). *Statistical Tables for Biological, Agricultural and Medical Research*. Edinburgh: Oliver & Boyd.

Gaito, J. (1965). Unequal intervals and unequal *n* in trend analyses. *Psychol. Bull.* **63**, 125–127.

Gale, A., Coles, M. & Blaydon, J. (1969*a*). Extraversion–introversion and the EEG. *Br. J. Psychol.* **60**, 209–223.

Gale, A., Dunkin, A. & Coles, M. (1969*b*). Variation in visual input and the occipital EEG. *Psychon. Sci.* **14**, 262–263.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Still, A. W. (1967). Use of orthogonal polynomials with nonparametric tests. *Psychol. Bull.* **68**, 327–329.

(*Manuscript received* 1 *August* 1970; *revised manuscript received* 28 *August* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 533–542
Printed in Great Britain

EXTRAVERSION–INTROVERSION, NEUROTICISM AND THE EEG: BASAL AND RESPONSE MEASURES DURING HABITUATION OF THE ORIENTING RESPONSE

BY ANTHONY GALE, MICHAEL COLES, PAUL KLINE AND VALERIE PENFOLD

Department of Psychology, University of Exeter

Occipital EEG was monitored during habituation trials to a regularly presented auditory stimulus. The subjects were selected in accordance with a factorial design which varied extraversion–introversion and neuroticism independently. A double-blind procedure was adopted for recording, scoring and statistical analysis. The EEG of extraverts and introverts was not differentiable; however, high neurotic groups were shown to have significantly greater EEG abundance than low neurotic groups across several frequency bands. There were no readily explicable interactions between the personality variables. For all subjects, two distinct response patterns appeared: (i) EEG attenuation (trials 1–3) and (ii) EEG augmentation (trials 4–20). Recovery time from stimulation may be longer than 10 sec. Both attenuation and augmenting responses are subject to the law of initial values. There were no differences in response amplitude or in speed of habituation for the personality groups.

The psychophysiology of individual differences is replete with first-rate questions and second-rate answers. It is reasonable to ask whether different personality groups differ also in terms of physiological reactivity. Unfortunately, this question is rarely asked properly. Studies of the relation between EEG and personality are no exception. As we have pointed out elsewhere (Gale *et al.*, 1969; Gale & Coles, 1969), experimental error of several types could account for the equivocal nature of the results. All those who have reviewed the field agree on this point (Claridge & Herrington, 1963; Claridge, 1967; Duffy, 1962; Eysenck, 1967, 1970; Fink, 1969; Gale *et al.*, 1969; Glaser, 1963; Goldstein & Sugerman, 1969; Harding, 1968; Hill, 1963; Mundy-Castle, 1955; Ostow, 1950; Sisson & Ellingson, 1955; Walters, 1964). Inspection of these reviews provides the following list of possible sources of experimental error: population sampling; unreliable and unvalidated measures of personality; crude clinical classification of subjects; failure to control for drug medication in chemotherapy; inattention to metabolic factors such as time since last meal and time of day; visual analysis of EEG; absence of measures of within or between marker reliability; non-random sampling of EEG; unjustified rejection of EEG records and/or subjects as 'unscorable'; misinterpretation of low-frequency analysis data; confusion about the relationship between frequency and amplitude; over-emphasis on alpha frequencies as opposed to the other frequency bands; non-standard recording procedures; possible experimenter bias during recording and scoring; effects of institutionalization on clinical subjects; implicit and explicit set within the subject; inappropriate statistical analysis; neglect of law of initial values; and finally, failure to provide a theoretical rationale for prediction of experimental outcome and consequent dependence on *post hoc* hypotheses. 'On-off' experiments are the general rule and there is a notable absence of parametric studies and replications.

The present study attempts to overcome some of these difficulties as follows: (1) subjects are selected by means of a well-factored personality inventory – the

EPI (Eysenck & Eysenck, 1964); (ii) a 'double blind' is employed for recording, scoring and statistical analysis, (iii) all subject groups undergo identical treatment, (iv) analysis of EEG is automatic and objective (low frequency analyser); (v) both basal and response levels are measured, (vi) recordings are taken at two fixed times of day only, (vii) analyses of variance are employed to estimate the effects of personality factors and their interaction, (viii) the population and experimental design are such as to meet with criticisms of previous work, made by Eysenck (1967), namely that in previous studies it is difficult to know whether obtained effects are attributable to extraversion, neuroticism or an interaction between the two, and (ix) consideration is given to the problem of the law of initial values

Three predictions are made on the basis of Eysenck's theory of the neurological substrates of individual differences in terms of arousal (extraversion) and activation (neuroticism). (1) The resting EEG will not discriminate between criterion groups on the basis of neuroticism as such. However, extraverts will have higher abundance than introverts (e.g. Gale *et al.*, 1969). (2) Again, response measures will discriminate only for extraversion; introverts will be more responsive than extraverts, and will take longer to habituate. (3) Interactions between extraversion and neuroticism will not override the main effects for extraversion and may even accentuate them.

METHOD

Sample. The EPI (Form B) was administered to 131 first- or second-year male students at the University of Exeter. The mean scores and standard deviations for this group were extraversion (E): 11·89, 4·46 and neuroticism (N) 9·91, 4·71. Correlation between E and N was negative and insignificant ($r = -0{\cdot}09$). An experimental sample of 60 subjects was then derived, making up six equal groups as shown in Table 1. Selection was performed by the third author and the personality score details of all subjects were coded and retained until analysis was complete.

Recording and analysis of EEG

Silver–silver chloride loose pad electrodes held on by a rubber net were applied trans-occipitally (Cooper *et al.*, 1969, p. 78) for bipolar recording, resistance being 7 kΩ or less. Primary record was obtained by a San'ei PG 802 polygraph calibrated to give 24 mm for 100 mV (peak to trough). TC 0·3 sec. Low-frequency analysis was provided by a San'ei EA 201 Analyser equipped with 10 pure bandpass filters as follows: (1) 2·0–4·5 c/sec.; (2) 4·5–6·5; (3) 6·5–7·5; (4) 7·5–8·5; (5) 8·5–9·5; (6) 9·5–10·5; (7) 10·5–11·5; (8) 11·5–14·5; (9) 14·5–20·0 and (10) 8·0–13·0. These filters are flat-topped and steep-skirted, and were tested for linearity of output with oscillators set at the midpoint for each frequency. A 5 sec. epoch was used. Integrated output was recorded in histogram form beneath the primary record on millimetre recording paper run at 5 mm per sec. between stimuli and 10 mm per sec. during stimulation periods. The integrated output yielded a range of 0–65 mm, measurable to the nearest millimetre.

Experimental procedure

The subject attended at either 2 p.m. or 4 p.m. The third author varied subjects to balance for time effects. The subject had been instructed by letter not to drink alcohol in the preceding 24 hr., to have a normal night's sleep and eat his usual lunch. The subject was prepared for a full polygraphic recording (results for EDA, FPV, EKG and respiration are to be reported later). He was led to a soundproof cubicle (Industrial Acoustics 4042) and lay supine on a bed. The ambient lighting level was 8·5 LM/ft.². The instructions to the subject were: 'After a few minutes you will hear a series of identical tones. They will not be unpleasant and you are not required to respond to them in any way. Try to relax, keeping your eyes closed. Do not fall asleep. Do not move around. Do you understand the instructions?' The cubicle door was then

closed. The polygraph was then calibrated and recording began. After a pre-test period of 4 min., 20 auditory stimuli (1000 c/sec., 65 db, duration 5 sec.) were presented automatically at a constant rate of one per 120 sec. through a loudspeaker above the subject's feet. A post-test period of 4 min. completed the procedure. Control apparatus reset the analyser so that for each stimulus, analyser epoch and stimulus were exactly synchronous.

Table 1. *The sample*

Personality group	Mean and range for extraversion	Mean and range for neuroticism
High extraversion	17·6	14·2
High neuroticism (HE HN)	17–19	11–17
Mean extraversion	12·3	13·1
High neuroticism (ME HN)	8–15	11–17
Low extraversion	4·5	16·3
High neuroticism (LE HN)	1–7	11–20
High extraversion	18·8	6·7
Low neuroticism (HE LN)	17–22	3–9
Mean extraversion	12·8	6·5
Low neuroticism (ME LN)	8–16	4–9
Low extraversion	4·6	6·8
Low neuroticism (LE LN)	1–7	4–10

Data reduction

Basal measure. Five consecutive 5 sec. epochs were sampled at the midpoint between all stimuli (i.e. 50–75 sec. after stimulation) for filters 1–9 and their values summed for each filter. Pre- and post-test samples were taken in similar manner at 'dummy' inter-stimulation points.

Response measure. Four epochs were sampled for each stimulus. A 1 (5 sec. prior, to point of stimulation onset), A 2 (during and coincident with stimulation), A 3 (offset, to 5 sec. post) and A 4 (5 sec. post to 10 sec. post). All ten filters were sampled in this manner, and an identical procedure adopted for 'dummy' pre- and post-stimuli.

Statistical treatment and results

Basal measures. Results are shown in Fig. 1. These were analysed by means of several analyses of variance (Myers, 1966), followed by Scheffé tests of multiple comparisons where appropriate.*

Main analysis

The main effects treated in this analysis were extraversion, neuroticism, filters and trials. Only the following main effects were significant. Neuroticism ($F = 5{\cdot}36$; d.f. = 1, 54; $P < 0{\cdot}05$): high neurotics have greater abundance than low neurotics. Filters ($F = 16{\cdot}18$; d.f. = 8, 432; $P < 0{\cdot}01$): integrated output is differentially distributed over the filters, the range 8·5–11·5 c/sec. being greater than both 2·0–8·5 c/sec. and 11·5–20·0 c/sec. The interactions filter/time, neuroticism/extraversion/time and the quadruple interaction were all significant beyond the 1 per cent level.

The absence of a significant interaction between neuroticism and extraversion indicates that the difference between high and low neurotics holds for all levels of extraversion. Although there is no such interaction, the relation between high and

* Limitation of space precludes inclusion of all summary tables of the analyses of variance. For method of computation, see Gale *et al.* (1969).

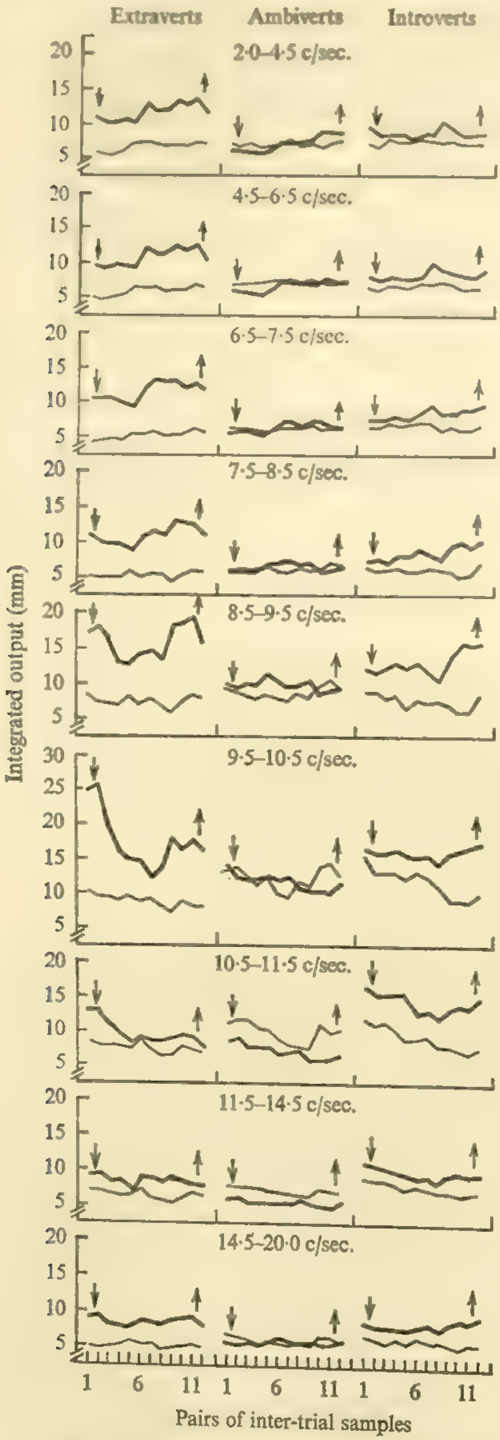


Fig. 1. Basal measures of EEG during habituation for each of the filters 1–9. Thick lines are high neurotic groups, thin lines low neurotic groups (10 subjects per group). Descending arrow: beginning of auditory stimulation. Ascending arrow: end of stimulus schedule. The abscissa represents a total recording time of 48 min. Each point plotted represents the mean value of two 25-sec. inter-trial samples.

low neuroticism is consistent only for the high and low extraversion groups. Analysis of the ambivert groups *alone* shows that in all cases the means for the high and low neuroticism groups are not distinguishable statistically. Thus the ambivert groups were omitted for the subsequent breakdown analyses, which are designed to clarify the neuroticism and interaction effects, as follows.

Breakdown analyses

Two sets of breakdown analysis were performed: (1) nine analyses for all filters taken separately, and (2) 12 analyses for pairs of adjacent trials. (Reference to Fig. 1 will enable the reader to follow the results of these analyses.)

(1) *Filters taken separately.* (*a*) The neuroticism effect holds beyond the 5 per cent level for all frequencies except 10·5–14·5 c/sec. (*b*) This difference is greatest for 7·5–9·5 c/sec., during the last third of the experiment ($P < 0{\cdot}01$). The difference is significantly reduced ($P < 0{\cdot}05$) during the middle period for 9·5–10·5 c/sec. (*c*) A triple interaction (neuroticism/extraversion/time) at 9·5–10·5 c/sec. shows that the main neuroticism effect is strongest for introverts in the later stages of the experiment, but strongest for extraverts at both the beginning and the end. Since this triple interaction ($P < 0{\cdot}05$) holds only at this frequency, it constitutes a possible source for the quadruple interaction in the main analysis. (*d*) Except for 14·5–20·0 c/sec., there is a time effect on all filters ($P < 0{\cdot}05$ at least in all cases). This shows that the range 2·0–8·5 c/sec. *increases* with time, and that 9·5–14·5 c/sec. *decreases* with time.

(2) *Pairs of trials taken separately.* (*a*) The neuroticism effect is located at both the onset and end of the experiment, but not in the middle period. (*b*) Neuroticism/filter interactions ($P < 0{\cdot}05$) show that the high/low neuroticism effect shifts gradually. At onset of stimulation it is located maximally at 9·5–10·5 c/sec. and by the final period it is maximal at 8·5–9·5 c/sec. (*c*) Finding 1 (*c*) is corroborated. Taken together, these two findings explain the interaction between filters and time in the main analysis.

Response measures

Separate analyses of variance were performed for all 10 filters. The main effects were extraversion, neuroticism, trials and analyser epoch. Since this analysis is concerned with *response* (i.e. changes across epochs A 1 to A 4), only main effects or interactions including A are considered. The following effects were significant.

(1) *Analyser epoch.* This holds at the 5 per cent level for 9·5–10·5 c/sec., and at 1 per cent for 2·0–6·5 and 10·5–20·0 c/sec. Also at 1 per cent for the broad-band alpha filter of 8·0 to 13·0 c/sec. In virtually all cases, integrated output *increases* during stimulation (i.e. during A 2) and *decreases* during A 3 and A 4. In Table 2 the epochs are shown for each filter, expressed in terms of an increase or decrease in relation to the previous epoch.

(2) However, significant interactions ($P < 0{\cdot}01$) between *epochs and trials,* on all filters except 6·5–7·5 c/sec., show that the *form* of response changes with trials. Thus (see Table 2) two stages are discernible: response type 1 and response type 2. Type 1 holds for the first three stimuli, where all filters show a *decrease* in output during stimulation (i.e. at A 2) with no consistent pattern of change for A 3. An exception to this is 2·0–4·5 c/sec., since it always shows a decrease at A 3. A 4 (5–10 sec.

Table 2. *Response data for all subjects combined, showing the change in response pattern during stimulation and the failure to habituate*

(A circle means that the value for this epoch is greater than that for the previous epoch. The overall pattern contributes to significant interactions between epochs and time. See Results section.)

Epoch ... Is epoch A2 greater than A1?

Filter ...	1	2	3	4	5	6	7	8	9	10
All trials	●	·	●	●	●	●	●	●	●	●
Pre-test	●	●	·	●	·	·	●	●	·	·
Trial										
1	·	·	·	·	·	·	·	·	·	·
2	·	·	·	·	·	·	·	·	·	·
3	·	·	·	·	·	·	·	·	·	·
4	·	·	●	·	·	·	·	●	●	·
5	●	·	●	●	·	●	●	●	●	●
6	●	●	●	·	●	●	●	●	●	●
7	·	·	●	●	●	●	●	●	●	●
8	·	·	●	●	●	●	●	●	●	●
9	·	·	·	·	●	●	●	●	●	●
10	·	·	●	●	●	●	●	●	●	●
11	·	·	·	●	●	●	●	●	●	●
12	·	●	·	●	●	●	●	●	●	●
13	●	·	●	●	●	●	●	●	●	●
14	●	●	●	●	●	●	●	●	●	●
15	●	●	●	●	●	●	●	●	●	●
16	●	●	●	●	●	●	●	●	●	●
17	●	·	●	●	●	●	●	●	●	●
18	·	·	·	●	●	●	●	●	●	●
19	●	●	●	●	●	●	●	●	●	●
20	●	●	●	●	●	●	●	●	●	●
Post-test	●	·	·	●	●	·	·	●	●	●
	2·0/4·5	4·5/6·5	6·5/7·5	7·5/8·5	8·5/9·5	9·5/10·5	10·5/11·5	11·5/14·5	14·5/20·0	8·0/13·0

Epoch ... Is epoch A3 greater than A2?

Filter ...	1	2	3	4	5	6	7	8	9	10
All trials	·	·	·	·	●	·	·	·	·	·
Pre-test	·	●	●	·	·	●	●	·	●	●
Trial										
1	·	·	·	●	·	●	●	●	●	●
2	·	●	·	●	●	●	·	·	●	●
3	·	●	·	●	●	●	●	●	●	●
4	·	●	●	●	●	●	●	·	●	●
5	·	●	·	●	●	·	●	·	·	●
6	·	●	·	●	●	·	·	·	·	·
7	·	·	·	·	·	·	·	·	·	·
8	·	·	·	●	·	·	·	·	·	·
9	·	●	·	●	●	·	·	·	·	·
10	·	·	·	·	·	·	·	·	·	·
11	·	·	·	·	·	·	·	·	·	·
12	·	●	·	·	●	●	·	·	·	●
13	·	·	·	·	·	●	·	·	·	·
14	·	·	·	·	·	·	·	·	·	·
15	·	·	·	·	●	·	·	·	·	·
16	·	·	·	·	·	·	·	·	·	·
17	·	·	·	·	·	·	·	·	·	·
18	·	·	●	●	●	●	·	·	·	●
19	·	●	●	●	·	●	●	·	·	●
20	·	·	·	·	·	·	·	·	·	·
Post-test	●	●	●	·	·	●	·	·	●	·

Epoch ... Is epoch A4 greater than A3?

Filter ...	1	2	3	4	5	6	7	8	9	10
All trials	●	·	·		·	·				·
Pre-test	●	·	●	·	●	●	·	●		●
Trial										
1	●	●	●	●	●	●	●	●	●	●
2	·	●	●	●	●	●	●	●	●	●
3	●	·	●	●	●				●	●
4	·		●	·	●		●			●
5	●	·	·	·	●	●	·			·
6	●	·				·	·			·
7	●	·	·	·			·			·
8	●	●	●							·
9	·	●	·		·				·	·
10	●	●	·		●			·	·	·
11	●	●	●			·			·	·
12	·	·	●		·	·				·
13	●				●					·
14	●	●	·							·
15	·	·	·		●					·
16	●	·	●				·			·
17	●	●			·	●	●			·
18	●	·				·	·			·
19	●	·	·		·	·	●			·
20	●	●		●	●		●		·	●
Post-test			●			●	●		●	●

after termination of stimulation) shows a tendency to increase over A3. *Thus type 1 shows a classic pattern of EEG attenuation. Type 2*, on the other hand, is *an EEG augmenting response*. Following trial 3, output *increases* during stimulation (A2), there is then a decrease which persists for A3 and A4. The transition from type 1 to type 2 occurs first on the higher-frequency filters.

(3) There are neuroticism analyser epoch effects ($P < 0{\cdot}05$) at 4·5–6·5 and 8·5–9·5 c/sec. (In the first case, stimulation increases output for both high and low neurotics, but more for the low neurotics, in the second case, stimulation increases output for the low neurotics only.)

(4) The only interaction between extraversion analyser epoch trials is at 2·0–4·5 c/sec. ($P < 0{\cdot}05$). However, the source of this interaction could not be located as a difference between extraversion groups in terms of point of transition from type 1 response to type 2.

Response measures and the law of initial values

Analyses here were performed on the alpha frequency broad-band filter only (8·0–13·0 c/sec.).

(1) *Law of initial values and type 1 and type 2 responses.* Two separate regressional analyses were performed for trial 1 (attenuation) and trial 10 (augmenting) for all subjects combined. The regression equation of Y (stimulus epoch A2) on X (pre-stimulus epoch, A1) was computed (Snedecor, 1956). In both cases the law of initial values is shown to operate. For trial 1 the regression equation is $Y = 0{\cdot}72X - 0{\cdot}43$: significance of slope: $t = 12{\cdot}00$, $P < 0{\cdot}001$; confidence limits $0{\cdot}86 > b > 0{\cdot}58$. For trial 10 the regression equation is $Y = 0{\cdot}73X + 9{\cdot}29$; significance of slope: $t = 10{\cdot}29$, $P < 0{\cdot}001$; confidence limits $0{\cdot}89 > b > 0{\cdot}57$.

Thus, in the case of response type 1, the greater the abundance of alpha activity prior to stimulation, the greater the attenuation. For response type 2, the lower the abundance of alpha activity prior to stimulation, the greater the augmentation. It is clear from the analyses of basal measures (see above) that type 2 occurs against a background of increased drowsiness.

(2) *Law of initial values and extraversion.* Since basal levels for extraversion are statistically indistinguishable, between-group comparison for extraversion, 'undoing' the law of initial values by means of analysis of covariance (Benjamin, 1963, 1967), is unnecessary.

(3) *Law of initial values and neuroticism.* In the case of neuroticism, however, basal differences *have* been established. Thus analyses of covariance were performed on trials 1 and 10 to examine whether removal of dependence of response amplitude on pre-stimulus level would have the effect of discriminating between high and low neurotics. However, the results of these analyses were negative.

Discussion

There are six principal findings in this study.

(1) The difference between extraverts and introverts reported in earlier work by ourselves and other authors is not replicated. Other authors have found that extraverts have higher abundance (Claridge & Herrington, 1963; Gale *et al.*, 1969;

Gottlober, 1938; Hume, 1968; Marton & Urban, 1966; Mundy-Castle, 1955; Savage, 1964). However, there are several studies with null or reverse results, where no difference holds or in which introverts have a higher alpha index (Broadhurst & Glass, 1969; Costa *et al.*, 1965; Fenton & Scotton, 1967; Glass & Broadhurst, 1966; Henry & Knott, 1941).

(2) There is a significant effect for neuroticism which is contrary to many previous findings and to our own expectations. In previous studies, if any effect is obtained, anxiety groups (equivalent to our HN LE group – see Table 1) have lower alpha abundance, higher high alpha and or high beta. Eysenck (1967) gives an extensive review of these findings. Savage (1964), using a design similar to ours, obtained no effect for neuroticism. Paradoxically, what Eysenck predicts for extraverts as opposed to introverts is what we obtain for high neurotics as opposed to low neurotics, and the null difference predicted for neuroticism holds for extraversion.

(3) There are no significant interactions between extraversion and neuroticism. This contradicts Savage's finding that HE HN have lower alpha abundance than HE LN; also, Eysenck's prediction that LE HN have fast EEG whereas HE HN have 'exceptionally slow EEG activity' (1967, p. 68). Fig. 1 indicates a possible interaction. The relation between HN and LN appears to be far more consistent for the HE and LE groups. However, the NE interaction is *not* significant and we must conclude that *all* E groups contribute to the main neuroticism effect.

(4) There are no psychologically interesting personality differences for habituation or for response amplitude. This confirms Fenton & Scotton's negative finding for extraversion (1967).

(5) The notion of EEG habituation itself appears to be suspect in this context. Indeed, it is difficult to know what is *meant* by 'habituation' in this study, since we have identified *two* types of response which follow one another closely in time: type 1, EEG attenuation; and type 2, EEG augmentation. Sokolov (1963) says that the OR returns during drowsiness, but he presents no detailed data. So far as we are aware, no other study has employed our particular measure of response. Tizard (1966) and McDonald *et al.* (1964) also fail to obtain EEG habituation in drowsy subjects by their method of response measurement. In the present study, two features of response are noteworthy. Type 2 follows type 1 almost immediately and both type 1 and, to a lesser extent, type 2 hold for *all* measured frequencies.

It may be the case that other investigators have failed to report the augmenting response simply because they have not looked for it. For example, visual analysis of EEG response with two consecutive trials without *attenuation* as the criterion for EEG habituation would lead to a *neglect* of post-criterion trials where augmenting does in fact occur. The existence of an augmenting response might well constitute an embarrassment for neurological models of habituation.

(6) Considerable attention has been paid to the law of initial values (LIV) in relation to electrodermal response (e.g. see Lacey, 1956; Hord *et al.*, 1964). So far as we are aware the problem of LIV has been virtually ignored in EEG studies. Since both attenuation and augmenting responses show a dependence of response amplitude on pre-stimulus level, future studies would clearly benefit from attention to this problem. We are conducting a parametric study designed to explore the phenomenon further.

We shall attempt briefly to explain the discrepancies between this study and previous studies. However, the full answer can only be provided by a series of thoroughgoing empirical, parametric studies, with control for many more of the variables mentioned in the introduction than we have achieved here. For example, our own previous study on extraversion and EEG differed from the present one in four ways: (*a*) the previous population was largely female; (*b*) recording was entirely in the morning; (*c*) the experimenters knew the subjects' personality scores prior to recording; and (*d*) the procedure and instructions were designed to keep the subject constantly *aroused*. Indeed, findings 1(*d*) and 2(*c*) above (see Results section), which indicate an increase in low frequency and a decrease in high frequency over time, show that the subjects in the present study became progressively more drowsy. The finding for neuroticism is unexpected yet hardly surprising in the light of criticisms of previous work (Claridge, 1967; Ellingson, 1954; Hill, 1963). These authors point out that the general finding of low-amplitude and high-frequency activity in anxiety states *may* be attributable to the tension induced by the experimental procedure. Apart from the fact that we employed a non-clinical population, our procedure is one reminiscent of techniques of relaxation therapy. It should be noted, however, that if the procedure *did* relax the high neurotic groups to an extreme, then the *instructions* themselves were enough to achieve this, since the relation holds *prior* to the habituation trials. Again a study varying neuroticism and instructional set is called for. The present study is of course, strictly speaking, only comparable to that of Savage, since extraversion and neuroticism are treated as *independent orthogonal factors*, whereas anxiety groups used in all other studies (apart from Savage, 1964) would be considered by Eysenck to be both introverted *and* neurotic (LE HN).

Finally, it may be argued that our long and constant inter-stimulus interval would be unlikely to produce habituation, or might even produce a conditioning to a time interval. In view of the complexity of the experiment in which five physiological measures were taken simultaneously, a regular interval was essential (*a*) for ease of administration and (*b*) to enable meaningful basal sampling and between-measure comparisons (in preparation). But the *duration* of type 1 and type 2 responses, as shown in Fig. 1, would set limits on the minimum inter-stimulus interval. Nor can it be argued that the technique of employing a 5 sec. epoch was too insensitive to detect changes in this experimental situation. The LIV analyses demonstrate the power of the measure. Indeed, 'response' may be even longer in duration than our time sample shows. The absence of a clear response pattern on the post-test 'dummy' stimulus (see Table 2) shows no evidence of temporal conditioning. Indeed, Table 2 provides rich new evidence of EEG response patterns. Again, a parametric study, varying regularity/irregularity and stimulus length, is called for.

In our view, a great deal more work must be done before definite conclusions may be reached concerning personality, basal EEG and EEG responsiveness. This study raises several problems concerning EEG measurement itself. It may well be a general rule that degree of experimental rigour is inversely related to the number of positive results.

This work was partially supported by a grant from the Medical Research Council (reference G969/185/C).

References

Benjamin, L. S. (1963). Statistical treatment of the law of initial values (LIV) in autonomic research: a review and recommendation. *Psychosom. Med.* 25, 556–566.

Benjamin, L. S. (1967). Facts and artifacts in using analysis of covariance to "undo" the law of initial values. *Psychophysiol.* 4, 187–206.

Broadhurst, A. & Glass, A. (1969). Relationship of personality measures to the alpha rhythm of the electroencephalogram. *Br. J. Psychiat.* 115, 199–204.

Claridge, G. S. (1967). *Personality and Arousal.* Oxford: Pergamon Press.

Claridge, G. S. & Herrington, R. N. (1963). An EEG correlate of the Archimedes spiral after-effect and its relationship with personality. *Behav. Res. Ther.* 1, 217–229.

Cooper, R., Osselton, J. W. & Shaw, J. C. (1969). *EEG Technology.* London: Butterworth.

Costa, L. D., Cox, M. & Katzman, R. (1965). Relationship between MMPI variables and percentage and amplitude of EEG alpha activity. *J. consult. Psychol.* 29, 90.

Duffy, E. (1962). *Activation and Behaviour.* New York: Wiley.

Ellingson, R. J. (1954). The incidence of EEG abnormality among patients with mental disorders of apparently non-organic origin: a critical review. *Am. J. Psychiat.* 111, 263–275.

Eysenck, H. J. (1967). *The Biological Basis of Personality.* Springfield: Thomas.

Eysenck, H. J. (1970). *The Structure of Human Personality.* London: Methuen.

Eysenck, H. J. & Eysenck, S. B. G. (1964). *Eysenck Personality Inventory.* London: University of London Press.

Fenton, G. W. & Scotton, L. (1967). Personality and the alpha rhythm. *Br. J. Psychiat.* 113, 1283–1289.

Fink, M. (1969). EEG and human psychopharmacology. *Ann. Rev. Pharmacol.* 9, 241–258.

Gale, A. & Coles, M. (1969). Brain waves and personality. *New Scientist* 43, 17–19.

Gale, A., Coles, M. & Blaydon, J. (1969). Extraversion–introversion and the EEG. *Br. J. Psychol.* 60, 209–223.

Glaser, G. H. (1963). The normal electroencephalogram and its reactivity. In G. H. Glaser (ed.), *EEG and Behaviour.* New York: Basic Books.

Glass, A. & Broadhurst, A. (1966). Relationship between EEG as a measure of cortical activity and personality measures. *Electroenceph. clin. Neurophysiol.* 21, 309.

Goldstein, L. & Sugerman, A. A. (1969). EEG correlates of psychopathology. In J. Zubin & C. Shagass (eds.), *Neurobiological Aspects of Psychopathology.* New York: Grune & Stratton.

Gottlober, A. B. (1938). The relationship between brain potentials and personality. *J. exp. Psychol.* 22, 67–74.

Harding, G. F. A. (1968). The psychological significance of the electroencephalogram. (Ap. Psych. Rep. 26. University of Birmingham at Aston.)

Henry, C. E. & Knott, J. R. (1941). A note on the relationship between personality and the alpha rhythm of the electroencephalogram. *J. exp. Psychol.* 28, 362–366.

Hill, D. (1963). The EEG in psychiatry. In J. D. N. Hill & G. Parr (eds.), *Electroencephalography.* London: Macdonald.

Hord, D. J., Johnson, L. C. & Lubin, A. (1964). Differential effect of the law of initial value (LIV) on autonomic variables. *Psychophysiol.* 1, 79–87.

Hume, W. I. (1968). The dimensions of central nervous arousal. *Bull. Br. psychol. Soc.* 21, 111. (Abstract.)

Knott, J. T. (1941). Electroencephalography and physiological psychology: evaluation and statement of the problem. *Psychol. Bull.* 38, 944–975.

Lacey, J. I. (1956). The evaluation of autonomic responses: toward a general solution. *Ann. N.Y. Acad. Sci.* 67, 123–163.

McDonald, D. G., Johnson, L. C. & Hord, D. J. (1964). Habituation of the orienting response in alert and drowsy subjects. *Psychophysiol.* 1, 163–173.

Marton, M. & Urban, Ya. (1966). An electroencephalographic investigation of individual differences in conditioning. (In Symposium on physiological bases of individual psychological differences.) *Proc. 18th Int. Congr. Exp. Psychol.*, I, Moscow.

Mundy-Castle, A. C. (1955). The relationship between primary-secondary function and the alpha rhythm of the electroencephalogram. *J. Nat. Inst. Personnel Res.* 6, 95–102.

Myers, J. L. (1966) *Fundamentals of Experimental Design.* Boston: Allyn & Bacon.

Ostow, M. (1950) Psychic function and the electroencephalogram. *Archs Neurol. Psychiat.* **64**, 385–400.

Savage, R. D. (1964) Electro-cerebral activity, extraversion and neuroticism. *Br. J. Psychiat.* **110**, 98–100.

[illegible], B. D. & Ellingson, R. J. (1955) On the relationship between normal EEG patterns and personality variables. *J. nerv. ment. Dis.* **121**, 353–358.

Snedecor, G. W. (1956) *Statistical Methods.* Ames, Iowa: Iowa State University Press.

Sokolov, E. N. (1963) *Perception and the Conditioned Reflex.* Oxford: Pergamon Press.

Tizard, B. (1966) Evoked changes in EEG and electrodermal activity during the waking and sleeping states. *Electroenceph. clin. Neurophysiol.* **20**, 122–128.

Walters, C. (1964) Clinical and experimental relationships of EEG to psychomotor and personality measures. *J. clin. Psychol.* **20**, 81–91.

(*Manuscript received* 21 *September* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 545–550
Printed in Great Britain

DIVERGENT THINKING: A COMPLEX FUNCTION OF INTERACTING DIMENSIONS OF EXTRAVERSION-INTROVERSION AND NEUROTICISM STABILITY*

By WILLIAM J. DI SCIPIO
Bronx State Hospital, New York, U.S.A.

Divergent thinking was measured by word fluency and word originality in a sample of 300 university students. After controlling for verbal intelligence, divergent thinking was found to be a complex function of extraversion introversion and neuroticism stability. Stable extraverts were significantly more fluent than stable introverts, but neurotic extraverts and introverts attained fluency scores approximating the mean of the extreme stable counterparts. The effect is explained in terms of Eysenck's arousal theory of personality types.

Introduction: a brief review of the literature

Research into the measurement of divergent thinking processes can be traced back to the early psychometric studies concerned with the faculty of 'imagination'. These studies demonstrated two distinct modes of expression: verbal fluency and originality. Verbal fluency defined a quantitative estimate of the rate of verbal output, while originality defined an estimate of the frequency of rare verbal responses with respect to a reference population. Reviews of the early work in this field may be found in Rogers (1953), Rim (1953), Eysenck (1960), Burt (1962, 1967) and Torrance (1967).

General intelligence, as defined by the early mental testers in Britain and America, was distinguished from the concept of 'imagination' in that general intelligence reflected the ability to reproduce information according to predetermined 'right' and 'wrong' test response criteria. Spearman (1927), however, held the view that tests of imagination scored for originality of response defined a specific cognitive factor which was only one of many others and might be employed in a wide battery of tests to yield a comprehensive estimate of the general intelligence factor g.

Hargreave's (1927) experimental studies of verbal fluency upheld Spearman's views and demonstrated that fluency, as measured by open-ended verbal tests, is a composite function of factors of general verbal ability, memory and motor speed of writing or speaking. Partial correlations among the fluency tests proved high enough to suggest that some further element remained to be accounted for (factor x) after removing the influence of the above factors. Hargreaves concluded that factor x was probably related to the absence of inhibitions, possibly a lack of self-criticism, thus attributing a conative explanation to verbal fluency.

Cattell (1934) placed greater emphasis upon the conative correlate of verbal fluency and renamed the tests of imagination 'Tests of Temperament'. In developing a battery of tests of perseveration and verbal fluency, Cattell found that a subtest of

* This paper formed part of a Ph.D. thesis by the author at the University of London, 1968.

the latter factor, Speed of Cognitive Output (SCO), proved to be the best predictor of temperament. The test required the subject to list all the two-syllable words which he could think of in 2½ min. The test correlated 0·30 ($n = 62$) with 'surgency', a personality trait denoting extraversion and measured by means of a self-rated questionnaire. The retest reliability of the SCO was 0·57 after one week.

The research to follow Cattell's lead continued to demonstrate the positive relationship of new verbal fluency tests to tests of temperament (Studman, 1935; Carroll, 1941; Notcutt, 1943; Gewirtz, 1948; Rim, 1954; Rogers, 1956; Hofstaetter *et al.*, 1957; Getzels & Jackson, 1962; Wallach & Kogan, 1965; Ramsey, 1966). Cattell's description of the surgency correlate of fluency as being one of sociability and the early replications of his findings led Eysenck (1960) to conclude that verbal fluency was related to extraversion, a personality type which incorporated traits of sociability and impulsivity.

Guilford (1950) introduced a new direction in the use of fluency and other divergent thinking tests by relating the abilities tapped by these tests to the abilities and patterns of character traits possessed by the 'creative' individual. The pattern of creative abilities was postulated to vary with the sphere of creative activity. Creative products varied with interests, attitudes and temperamental factors. Guilford's approach was useful, but only within the context of his factor structure of the intellect. Divergent thinking tests became loosely defined criteria of 'creativity', which did not account for the social acceptance of creative output for a given field.

Although Guilford's influence diverted further investigations from Cattell's and Eysenck's notion of fluency as a function of extraversion, the evidence continued to suggest that the individual who scored highly on fluency or 'creativity' tests tended to be impulsive, sociable, talkative and uninhibited, in addition to having a wide range of interests and preferring arts to sciences (Barron, 1953; Getzels & Jackson, 1962; Hudson, 1966; Bowers, 1967; Schaefer & Anastasi, 1968).

White (1968), in a recent paper, reports on the distinction between the personality correlates of creative behaviour as determined by known groups of creative individuals in contrast to creative individuals determined by high scores on divergent thinking tests. He suggests that personality correlates of divergent thinking tests are not *a priori* measures of creative ability as displayed by individuals who are judged by others as possessing 'creative talent'. White presents evidence for 200 male university freshmen which indicated that the extravert, as defined by Cattell's 16 PF, obtains higher scores on divergent test measures (flexibility, fluency and originality) than the introvert, with the relationship applying in the same order for stable *v.* neurotic personalities.

A contradictory note on the subject of personality correlates of divergent thinking is introduced by Rim (1953) in a study which attempted to establish a differential diagnostic criterion for hysterics (neurotic extraverts) and dysthymics (neurotic introverts). Rim did not find a significant difference between these clinical groups in verbal fluency.

It appears therefore that the only firm conclusion to be drawn from the literature at present is that divergent thinking varies as a function of personality typology in normal individuals. Several aspects of intellectual capacity and sex differences which may be relevant in this area have not been accounted for in the design of many

studies. The need for further research is therefore obvious. The work reported in this paper tests the hypothesis that divergent thinking, as measured by word fluency and word originality, is a function of personality variables of extraversion–introversion and neuroticism–stability when the influence of general verbal intelligence has been accounted for. Extraverts are predicted to be more fluent and original than introverts. Measures of the relationship of neuroticism–stability to divergent thinking are included in order to assess further possible interactions of personality types in modes of thinking.

Method

Description of the sample. A total of 100 male and 200 female students was tested at a large American university with a student population of 20,000. An approximate cross-sectional sample was obtained by setting age limits of 16–25 years (mean age = 19 yr.) and recruiting both voluntary students and others required to take experimental examinations as part of their introductory psychology courses. The students were tested by the same examiner in groups varying in size from six to 61. Testing atmosphere was similar to most other examinations familiar to the students in the university setting. The general purpose of the tests was described as experimental and in no way relevant to the students' university academic status. Twenty students volunteered to return for results and were retested for reliability estimates.

Description of the test measures. The assessment of general intellectual level was made by the Scholastic Aptitude Test – Verbal (SAT-V), which was administered approximately 4–6 months prior to admission by the university as a required entrance examination. The scores were available for 53 of the male and 134 of the female sample.

The experimental test session consisted of the following measures: (*a*) Eysenck Personality Inventory, Form A (EPI), a self-rated 57-item questionnaire designed to measure the independent dimensions of extraversion–introversion and neuroticism–stability; (*b*) Speed of Cognitive Output Test (SCO; Cattell, 1934), which requires the subject to list as many two-syllable words as he can think of, with the limitations that they not be objects in the room. Subjects were asked to circle the last word completed at a signal which was given at 30 sec. intervals. Verbal fluency scores were based upon the total number of words for a 2½ min. period.

Word originality scores were obtained by assigning a score of one to each word used by a subject which did not appear in a reference sample of 100 subjects (33 male, 67 female) selected randomly from the full sample pool of 300 subjects.

Analysis of data. Dichotomous groups of introverts and extraverts were determined by dividing the total sample at the mean EPI score for extraversion–introversion. The same procedure was applied to form dichotomous neurotic and stable groupings. Mean verbal fluency and originality scores were computed for each independent sex and personality variable. In order to minimize loss of subjects in an attempt to maintain proportional cells for analysis of variance, three factorial designs were computed for each main effect and interaction of interest. An analysis of covariance employing the SAT-V as a single covariate was included in the design, although further loss of subjects whose SAT-V scores were not available was unavoidable. An additional covariance analysis was included which accounted for verbal fluency scores as a covariate when taking originality scores as the dependent variable.

Results

There were no significant differences between males and females for either personality (EPI) or verbal intelligence (SAT-V) variables. The orthogonality of the E–I and N–S personality dimensions was upheld ($r = -0{\cdot}17$, $n = 300$).

Word fluency. Table 1 presents means, S.D.s and numbers for the sample with SAT-V scores available for analysis. The results of the 2×2 analyses of variance for this sample contain two significant personality effects: extraverts are significantly more fluent than introverts ($F = 3{\cdot}95$; d.f. = 1, 154; $P < 0{\cdot}05$), and a significant

interaction of extraversion–introversion and neuroticism–stability ($F = 8{\cdot}94$; d.f. = 1, 154; $P < 0{\cdot}003$). The interaction effect qualifies the main effect in that stable extraverts are significantly more fluent than stable introverts. The neurotic extravert is *less* fluent than the stable extravert and the neurotic introvert is *more* fluent than the stable introvert. Both neurotic groups approximate a mean fluency value lying between their stable counterparts. The covariance analyses, which covaried SAT-V scores, revealed that the above personality differences increased in level of significance ($F = 5{\cdot}83$; d.f. = 1, 153; $P < 0{\cdot}02$; and $F = 9{\cdot}50$; d.f. = 1, 153; $P < 0{\cdot}002$ respectively). Females were significantly more fluent than males ($F = 4{\cdot}46$; d.f. = 1, 280; $P < 0{\cdot}05$). Sex differences in fluency scores did not emerge as significant effects after covariance analysis.

Table 1. *Means and S.D.s of word fluency scores for personality dimensions summed over sexes*

	n	Mean	S.D.
Neurotic extraverts	45	25·49	7·65
Neurotic introverts	45	26·24	8·03
Stable extraverts	34	29·29	7·37
Stable introverts	34	22·59	7·94

Word originality. Word fluency and originality scores of the SCO correlated 0·67 ($n = 100$) for males and 0·61 ($n = 200$) for females. All analyses applied to the verbal fluency scores were repeated on originality scores. The single significant effect to emerge was that males were significantly more original than females ($F = 6{\cdot}10$; d.f. = 1, 280; $P < 0{\cdot}02$). This finding was upheld in covariance analyses when covarying both SAT-V scores and verbal-fluency scores ($F = 7{\cdot}20$, d.f. = 1, 181, $P < 0{\cdot}01$; $F = 12{\cdot}32$, d.f. = 1, 181, $P < 0{\cdot}001$).

Test–retest reliability coefficients were 0·92 for word fluency and 0·64 for word originality for 20 subjects tested at approximately one-week intervals.

DISCUSSION

The SCO test appears to be a better correlate of general verbal ability (SAT-V) for females ($r = 0{\cdot}30$; $n = 134$) than for males ($r = 0{\cdot}04$; $n = 53$) when scored for word fluency as well as for word originality ($r = 0{\cdot}29$ and $0{\cdot}03$ respectively). After removing the influence of the fluency score from the originality score, there is less evidence of a sex difference for the relationship of originality to verbal intelligence ($r = 0{\cdot}17$ and $0{\cdot}08$ respectively). The low positive correlation between word fluency and verbal intelligence for females is consistent with the findings first reported by Hargreaves (1927), but is now qualified by the sex difference in the present sample. The originality variable, however, shows very little relationship to intelligence for both sexes, which is in unqualified agreement with Hargreaves. A possible explanation of this finding may concern the operational criterion of word originality used in this study. Originality, as defined by unusual or rare two-syllable word production, does not involve attention to verbal meaning or verbal context and therefore does not tap vocabulary ability or general verbal reasoning. The word fluency measure,

however, is closely related to a general store of vocabulary and consequently this variable may be expected to correlate with the SAT-V test because of the shared variance both tests contain for vocabulary ability.

A further result which lends support to the independence of scores of fluency and originality within the same test is the significant sex differences found in opposite directions for each variable. Females are found to be more fluent than males. On the other hand, males are found to give more original responses than females. These sex differences are somewhat equivocal with respect to covariance analysis in that verbal IQ, the covariate, does not appear to be a linear function of word fluency for males, nor word originality for both sexes. The analysis of personality differences is therefore combined for both sexes because of the statistical advantage of increased sample size and the positive linear relationship between verbal IQ and word fluency for the full sample ($r = 0{\cdot}20$; $n = 300$).

The most salient finding to appear in this study is a qualification of the stated hypothesis when considering the interaction of the neuroticism–stability dimension. The hypothesis that divergent thinking as measured by word fluency and originality is higher for extraverts than introverts was upheld for word fluency but not originality. Neuroticism is found to decrease the fluency of extraverts and increase the fluency of introverts. Both neurotic groups of E and I approach the overall mean of the stable E and I groups. This finding explains why Rim (1953) was not able to find a significant word fluency difference between his two criterion groups of hysterics and dysthymics. Although both of Rim's clinical groups represent extremes in E and I, both are clinical conditions of high neuroticism.

A general *post hoc* formulation to account for the E–N interaction and based upon neurotic instability may be formulated in the following terms. Traits such as verbal fluency are manifested in extreme forms in opposite directions for introverts and extraverts, if both groups are stable or non-neurotic. Defining neuroticism or high anxiety as a disturbance in normal function, the mechanism which would normally maintain traits at upper and lower limits can no longer be assumed to function at these limits consistently. Consequently, the behaviour would regress toward the mean of the stable groups. In accordance with Eysenck's (1967) theory of arousal and personality, the mechanism which underlies the control of traits such as verbal fluency may be postulated as one of cortical inhibition versus excitation. In its most general form, this theory states that extraverts function with minimal levels of cortical inhibition of the central nervous sytem, while introverts function with maximal levels. The extravert would therefore respond behaviourally in a more impulsive and divergent manner than the introvert. Neuroticism introduces a weakening of higher cortical functioning and consequently would bring about an inability to maintain personality traits usually manifested by the E–I continuum in stable individuals. Predictions could no longer be made in relation to high and low verbal fluency, except that overall performance of neurotics at either end of the extraversion–introversion continuum would fall somewhere between the extremes of the stable groups.

The study suggests implications for clinical research in so far as divergent thinking tests may prove to be a useful diagnostic instrument in assessing the influence of neuroticism or high anxiety on spontaneous verbal output. Changes during the course of therapy might be monitored by assessing the rate at which the individual

approaches a verbal fluency score appropriate to his sex, intellectual level and stable personality type.

I wish to thank Drs W. A. Winnick and G. A. Pierson, Queens College, New York, for administrative assistance, and Maria Di Scipio, R. N. and Dr M. M. Singh, Bronx State Hospital, New York, for reviewing the manuscript. The Ph.D. thesis was sponsored by Professor H. J. Eysenck, Institute of Psychiatry, London.

REFERENCES

BARRON, F. (1953). Complexity–simplicity as a personality dimension. *J. abnorm. soc. Psychol.* **48**, 163–172.

BOWERS, P. G. (1967). Effect of hypnosis and suggestions of reduced defensiveness on creativity test performance. *J. Personality* **35**, 311–322.

BURT, C. (1962). The psychology of creative ability. *Br. J. educ. Psychol.* **32**, 292–298.

BURT, C. (1967). Book review of *Contrary Imaginations* by Liam Hudson. *Br. J. Psychol.* **58**, 173–174.

CARROLL, J. B. (1941). A factor-analysis of verbal abilities. *Psychometrika* **6**, 279–308.

CATTELL, R. B. (1934). Temperament tests. II. Tests. *Br. J. Psychol.* **24**, 20–49.

EYSENCK, H. J. (1960). *The Structure of Human Personality.* London: Methuen.

EYSENCK, H. J. (1967). *The Biological Basis of Personality.* Springfield: Thomas.

GEWIRTZ, J. L. (1948). Studies in word-fluency. II. Its relation to eleven items of child behaviour. *J. genet. Psychol.* **72**, 177–184.

GETZELS, J. W. & JACKSON, P. W. (1962). *Creativity and Intelligence: Explorations with Gifted Students.* London: Wiley.

GUILFORD, J. P. (1950). Creativity. *Am. Psychol.* **5**, 444–454.

HARGREAVES, H. L. (1927). The faculty of imagination. *Br. J. Psychol. Monogr. Suppl.* **3**, 10.

HOFSTAETTER, P. R., O'CONNOR, J. P. & SUZIEDELIS, A. (1957). Sequences of restricted associative responses and their personality correlates. *J. gen. Psychol.* **57**, 219–227.

HUDSON, L. (1966). *Contrary Imaginations: a Psychological Study of the English Schoolboy.* London: Methuen.

NOTCUTT, B. (1943). Perseveration and fluency. *Br. J. Psychol.* **33**, 200–208.

RAMSAY, R. W. (1966). Speech patterns and personality. (Unpublished Ph.D. thesis, University of London.)

RIM, Y. (1953). Perseveration and fluency as measures of introversion–extraversion in abnormal subjects. (Unpublished Ph.D. thesis, University of London.)

RIM, Y. (1954). Perseveration and fluency as measures of introversion–extraversion in abnormal subjects. *J. Personality* **23**, 324–334.

ROGERS, C. A. (1953). The structure of verbal fluency. *Br. J. Psychol.* **44**, 368–380.

ROGERS, C. A. (1956). The orectic relations of verbal fluency. *Aust. J. Psychol.* **8**, 27–46.

SCHAEFER, C. E. & ANASTASI, A. (1968). A biographical inventory for identifying creativity in adolescent boys. *J. appl. Psychol.* **52**, 42–48.

SPEARMAN, C. (1927). *The Abilities of Man: their Nature and Measurement.* London: Macmillan.

STUDMAN, L. G. (1935). Studies in experimental psychiatry. V. 'W' and 'F' factors in relation to traits of personality. *J. ment. Sci.* **81**, 107–137.

TORRANCE, E. P. (1967). New types of items for measuring the creative thinking abilities. *International Workshop on Possibilities and Limitations of Educational Testing.* Berlin: Pädagogisches Zentrum.

WALLACH, M. A. & KOGAN, N. (1965). *Modes of Thinking in Young Children.* New York: Holt, Rinehart & Winston.

WHITE, K. (1968). Anxiety, extraversion–introversion, and divergent thinking ability. *J. creative Behav.* **2**, 119–127.

(*Manuscript received* 23 *June* 1970; *revised manuscript received* 7 *August* 1970)

Br. J. Psychol. (1971), 62, 4, *pp.* 551–555
Printed in Great Britain

THE PERFORMANCE OF APPROVED-SCHOOL BOYS ON THE GIBSON SPIRAL MAZE

By D. N. MARTIN and SUSAN WARDE
Red Bank School, Newton-le-Willows, Lancashire

Delinquents generally perform differently on the Gibson Spiral Maze from non-delinquents. However, they are tested under different circumstances and this may explain the difference. Evidence is presented here that performance on the maze is indeed influenced by the circumstances of its administration. New validity studies which control these influences are needed. In this study, comparisons of more and less delinquent boys tested similarly generally fail to support the claim that delinquency and psychomotor style (as exhibited by this test) are related.

The Spiral Maze (Gibson, 1964, 1965) is a simple paper-and-pencil psychomotor test which arose out of research with the Porteus Maze Test (cf. Porteus, 1965). It has been known for some time that delinquent populations obtain higher Q scores on the Porteus Mazes than non-delinquent groups. The Q score is a composite measure of some qualitative aspects of performance on the test, which is basically a measure of intelligence. Unfortunately the Porteus Mazes are lengthy to administer and the calculation of the Q score can be complicated. Furthermore, it represents not only psychomotor errors but also some cognitive ones. These difficulties led Gibson to design the Spiral Maze as a purer and more convenient measure of the psychomotor variables which appear to be implicated in delinquency.

It consists of a spiral pathway, boldly printed on a large card. The subject traces his way with a pencil from the centre of the maze to the exit, avoiding both the lines at the side and a series of small round obstacles. Every 15 sec. he is urged to go quickly. The psychologist records the *time* taken to complete the maze and calculates an *error* score from the number of contacts with lines and obstacles.

In most completed validity studies with the new test (Gibson, 1965; Orme, 1968) delinquents have performed, as predicted, more quickly and carelessly than their non-delinquent counterparts. However, they have always been tested under different circumstances (i.e. in institutions for offenders, and with court appearances behind them) and Gibson has pointed out that differences in performance could simply reflect these circumstances. This criticism (which can be levelled at much of the previous work on psychomotor style and delinquency) is not merely academic. Two studies (Gillon, 1965; Clarke, 1968) in which apparently representative groups of delinquents completed the maze slowly and carefully raise the possibility that situational variables may indeed influence performance. The first object of the present study was to confirm that this is the case, and comparisons were made between the performance of boys tested in different types of school and for the second rather than the first time. If it is the case, the validity of the test needs to be checked, and the second aim was to contribute some new validity data based on comparisons between more and less delinquent boys tested under the same circumstances. Since absconding is closely related to prognosis (Clarke & Martin, 1971) it was also expected that the absconders in a group would be relatively quick and careless in their performance.

Method

Following committal to approved school, boys spend a short period at a regional classifying school, where they are assessed, and are then moved to an appropriate training school. This is normally the school nearest their home.

Out of an original group of 75 consecutively admitted intermediate (13–15 years) approved school boys tested on the Spiral Maze in the classifying school, 47 were tested again in their training schools and constitute the retest sample. As a control sample, 86 of the 100 classifying school intermediates preceding the earliest member of the retest group were tested in their training schools only. Although the samples are not consecutive, they are probably representative. Cases were lost mainly because three training schools, which received few cases and are distant from the classifying school, were not visited. The same five training schools were visited for both samples. The retest and control samples were comparable in age at admission to the classifying school (retest $\bar{x}$ = 14 years 1 month; control $\bar{x}$ = 14 years 3 months), WISC or SB L-M IQ (retest $\bar{x}$ = 91·5; control $\bar{x}$ = 92·7), and average length of time between transfer from classifying school and training school test (retest $\bar{x}$ = 22 weeks; control $\bar{x}$ = 23 weeks). With the exception of 19 in the classifying school, all tests were carried out by the same psychologist.

In the case of the retest sample, three indices of delinquent behaviour, known to be predictive of reconviction (Home Office, 1969), were noted for each boy. These were: age at first court appearance, number of findings of guilt, and shortest gap between successive unrelated court appearances. It was also known, for boys in *both* samples, if an absconding from training school occurred within a follow-up period of 8 months to 1 year (most abscondings occur early in training).

Non-parametric statistical methods were used throughout the study since distributions of T (time) and E (error) scores were generally skewed. Significance tests were two-tailed.

Results

Situational influences on performance

Table 1 gives the median T (in sec.) and E scores obtained under different testing circumstances. The retest sample's training school performance was significantly quicker (Mann–Whitney $U = 520·5$; $P < 0·01$) but not significantly more careless ($U = 884·0$, n.s.) than its classifying school performance. The performance of the control sample, tested only in training schools, was significantly quicker ($U = 1537·5$; $P < 0·02$) than the retest sample's classifying school performance, although not significantly more careless ($U = 1851·0$, n.s.). It was significantly slower ($U = 1450·5$; $P < 0·01$) and more careful ($U = 1524·5$; $P < 0·02$) than the retest sample's training-school performance.

Table 1. *Median T and E scores for retest and control samples*

	Retest sample (n = 47), Classifying School	Retest sample (n = 47), Training School	Control sample (n = 86)
Median T (sec.)	45·0	35·0	40·0
Median E	10·0	13·0	11·0

It was also found, using Kruskal–Wallis one-way analyses of variance, that there were significant differences in training school T ($H = 12·10$; d.f. = 4; $P < 0·02$) and E ($H = 11·04$; d.f. = 4; $P < 0·05$) scores between the retest sample subjects tested at different training schools. When the classifying school scores of the same boys, grouped in the same way, were analysed in this manner, no such significant

differences were found (T: H = 7·19, d.f. = 4, n.s.; E: H = 1·54, d.f. = 4, n.s.). Significant differences were also found between the control sample boys allocated to different training schools.

These results raise the possibility that the retest training school sample as a whole may have performed more quickly and carelessly than the control sample as a whole, simply because a greater proportion of its members happened to have been allocated to the training schools which produced the hastiest performances. However, 32 per cent of the retest sample as opposed to 36 per cent of the control sample were tested in the two schools which produced the quickest performances for the respective samples, and 51 per cent as opposed to 45 per cent in the two slowest schools. A similar pattern obtained for E scores.

Three supplementary analyses were carried out. First, age at the time of testing was correlated (Spearman) with the Spiral Maze scores of the retest group in the classifying school and of a random sample of 30 from the control group. The correlation of age with T was in both cases +0·18 and no correlations were significant. Secondly, the length of follow-up between transfer and testing in training schools was correlated with T (rs = −0·13, n.s.) and E (rs = −0·05, n.s.). Finally, comparisons were made between the scores of the classifying school test group boys who were and were not tested by the psychologist who administered tests in the training schools. There were no significant differences in T (U = 264·5, n.s.) or E (U = 242·0, n.s.) scores.

Validity data

In this section Spiral Maze performance is considered in relation to absconding and indices of delinquent behaviour. Analyses were carried out separately for the retest sample in the classifying school and in the training school, and for the control sample (absconding data only). This was because, in view of the findings in the last section, the relative standing of the subjects' performance could change under different circumstances.

Throughout this section, T/E scatter plots were partitioned (i.e. into quick and careless, slow and careful zones, etc.) by regression lines as advocated by Gibson (1965).

The between-schools effects noted above could have vitiated any straight correlation of the scores of either of the training school samples with other data. For example, a real difference between absconders and non-absconders within each school separately could be masked, or a spurious difference between all absconders and all non-absconders in a sample could be produced by a disproportionate representation of absconders in, say, a school which produced very careless performances. Therefore each boy's training school T and E scores were converted to percentiles according to his standing in relation to members of the same sample tested in the same school. Natural scores were used in the case of the classifying school performance of the retest sample.

(*a*) *Absconding*. Comparisons (Mann–Whitney U test) on T and E separately were made between the absconders and non-absconders in each of the classifying school and two training school groups. In addition, the distributions of absconders and non-absconders among the four zones of the T/E scatter plots were examined visually and

by χ^2. Table 2 summarizes the data, which are in the form of median natural scores for the classifying school sample and mean within school percentiles for the training school samples. None of the analyses in relation to absconding gave a significant result.

Table 2. *T and E scores of absconders and non absconders in groups tested under different circumstances*

	Retest sample, Classifying School (medians)	Retest sample, Training School (mean percentiles)	Control sample (mean percentiles)
Time			
Absconders	44·5 (n = 25)	42·1 (n = 25)	43·9 (n = 48)
Non-absconders	46·0 (n = 22)	48·9 (n = 22)	46·7 (n = 38)
Error			
Absconders	8·0 (n = 25)	46·6 (n = 25)	42·7 (n = 48)
Non-absconders	11·0 (n = 22)	39·3 (n = 22)	47·7 (n = 38)

(*b*) *Indices of delinquent behaviour*. The three indices of severity of delinquent behaviour (age at first court appearance, number of findings of guilt, shortest gap between court appearances) were correlated with the T and E scores obtained by the retest sample separately in the classifying school and in the training school. Of the 12 (Spearman) correlations, only one was significant. The duration of the shortest gap between court appearances was related as predicted to error scores obtained by the retest group in the training school ($rs = -0·33$; $P < 0·05$). The delinquency index scores of boys falling in different zones of the $T.E$ scatterplots were compared by means of Kruskal–Wallis analyses of variance, but there were no significant effects.

DISCUSSION

The main conclusion is that the boys who were tested in the classifying school performed differently on the Spiral Maze from boys tested in training schools, and that training school boys who had done the test before performed differently from those who had not.

It is unlikely that boys in different groups differed in 'true' ability to traverse the maze. For various reasons the effect of having done the test before does not seem to have been a conventional practice effect, and the unlikely possibility that training school experience actually modifies psychomotor skills is not supported by the low correlations between follow-up period and scores on the maze. However, age is one variable on which groups were not strictly comparable, both training school groups being older than the classifying school group *at the time of testing*. There is some evidence (Gibson, 1969) that T decreases sharply with increasing age. In the present study, however, the relationship between age and Spiral Maze scores, over a 2- to 3-year range, was very slight and it seems unlikely that it could have influenced the results.

One possible explanation of the results concerns the manner of the test's administration. Different testers produced similar results in the classifying school. However, since the psychologist was aware of the general design of the study, the test may unwittingly have been administered in a different way in the classifying school, in training schools, and to boys who had been tested before. The correlation between T

and E which is held to reflect the degree of time-stress induced (Gibson, 1965), did vary between these testing conditions, increasing as T decreased. However, it should increase as T decreases *for whatever reason*, and the variation provides no more support for the possibility of differing testing techniques than it does for other interpretations of the results. Moreover, differences in the scores of boys sent to different training schools, which also suggest a relationship between performance and circumstances of testing, were presumably not a result of the tester's enthusiasm for the research since this analysis was not foreseen at the time of testing.

The differences between groups probably were the result of varying test taking expectations and attitudes prompted by the different circumstances under which they were tested. If performance on the Spiral Maze is very sensitive to these influences, the customary quick and careless performance of delinquents may not be a characteristic style but a reflexion of their particular attitudes. In this case, there is a need for new validity studies in which this possibility is controlled. Unfortunately, the present comparisons between more and less delinquent subjects tested similarly throw little light on the problem. Under different conditions, associations with absconding and criteria of delinquency failed to reach significance in all cases but one. This could be because the selected population of approved schools is relatively homogeneous in delinquent tendency and the results could be seen as a comment on the test's *utility* in this setting rather than on its validity in general. Gibson has undertaken a longitudinal study in which situational variables appear to be controlled, but complete results are not yet available (cf. Gibson, 1969).

The present study suggests that the evaluation of individual records should be based on detailed local norms.

We wish to thank the staff of the schools involved in this research for providing facilities, and Dr R. V. G. Clarke and Dr D. C. Kendrick for their very helpful suggestions.

References

Clarke, R. V. G. (1968). Factors influencing absconding by approved-school boys. (Unpublished Ph.D. thesis, University of London.)

Clarke, R. V. G. & Martin, D. N. (1971). *Absconding from Approved Schools*. (Home Office Research Studies.) London: H.M.S.O.

Gibson, H. B. (1964). The Spiral Maze: a psychomotor test with implications for the study of delinquency. *Br. J. Psychol.* **55**, 219–225.

Gibson, H. B. (1965). *Manual of the Gibson Spiral Maze*. London: University of London Press.

Gibson, H. B. (1969). The Gibson Spiral Maze test: retest data in relation to behavioural disturbance, personality and physical measures. *Br. J. Psychol.* **60**, 523–528.

Gillon, P. W. (1965). The assessment of impulsiveness in female delinquents. (Unpublished M.A. thesis, University of London.)

Home Office (1969). *The Sentence of the Court*. London: H.M.S.O.

Orme, J. E. (1968). Psychological functioning in girls at a classifying school. *Appr. School Gaz.* April, pp. 11–15.

Porteus, S. D. (1965). *Porteus Maze Tests: Fifty Years Application*. Palo Alto: Pacific Books.

(*Manuscript received* 17 *June* 1970; *revised manuscript received* 1 *October* 1970)

Br. J. Psychol. (1971), 62, 4, pp. 557–565
Printed in Great Britain

NEW TESTS OF MUSICAL APTITUDE

By J. B. DAVIES
Department of Psychology, University of Strathclyde, Glasgow

Existing tests of musical ability, with the exception of the Seashore battery (1919), have made extensive use of musical material. Since the musical experience of individuals may differ in extent and type, some experiential bias may result. In the present study, a new series of tests of musical aptitude is described which makes no use of formal musical material, but which involves non-musical or quasi-musical forms. Analysis of results obtained with the tests suggests that such an approach is feasible, and that the battery operates in a meaningful and consistent manner.

For measures of musical aptitudes to be most useful in the selection and measurement of natural musical talent, any culture-specific or experiential bias in the testing situation should be minimal. This requirement limits the usefulness of any testing system which employs material of a formally musical nature, because (*a*) differences in musical training of subjects, or in the nature of their past musical experience, result in differences in familiarity with, and ease of manipulation of, musical materials; (*b*) musical conventions vary between cultures, and between subcultures in the same society. Therefore any measures employing material of a formally musical nature may be contaminated by the musical environment and experience of subjects, which are probably not the best index of musical aptitude.

The phrase 'tests of musical ability', or similar, is used widely enough to create some confusion (Wing, 1941; Mainwaring, 1931; Bentley, 1966; Lundin, 1949). Ability involves factors of learning (including practice) and aptitude, which combine to produce a certain level of performance or attainment. Aptitude, however, does not necessarily manifest itself, and can theoretically exist in the absence of any performance, i.e. where there is no *apparent* ability.

The futility of devising a test to measure aptitude for learning French, in which all the items are written in French, is easily seen. However, with the notable exception of Seashore (1919, 1938) most authors have used material which is in essence musical, or even intact musical extracts and structures (Wing, 1948; Hevner & Landsbury, 1935; Lowery, 1926). The view held by Mursell (1937) that 'only the observations of the subjects in various musical situations are a guide to the degree to which talent is present' is not tenable as an *a priori* basis for test construction. Any situation can give a guide to the degree to which talent is present, if we can show that it meets certain criteria of validity and reliability. Anastasi (1961) writes: 'It should be noted in this connexion that the test items need not resemble closely the behaviour the test is to predict. It is only necessary that an empirical correspondence be demonstrated between the two.'

Aims of the investigation

The aims of the present study were (*a*) to assess the feasibility of constructing a battery of tests involving no material of a formally musical nature, and (*b*) to find out if such a test battery could be made to meet adequate criteria of validity and reliability.

In this study, the point made by Wing (1948), that music teachers are better disposed towards items of a 'musical' nature and therefore more likely to use them, was considered, but rejected as a yardstick by which to construct tests.

The study fell into two parts. The first involved an examination of all the types of material used by test constructors in existing tests of musical aptitude, and the selection of certain of these for further examination; the second involved the compilation of a test battery, and its progressive modification as a result of administration to over 2000 school children, aged between 7 and 11 years.

Preliminary considerations

As a result of the examination of (*a*) existing test batteries, (*b*) factorial studies of different testing systems, and (*c*) results of a questionnaire survey of tutors in music colleges, tests of pitch, interval, melody and rhythm were selected for further development. A detailed account of reasons for rejecting other types of test material is given in Davies (1969). Any material involving formal musical knowledge, notation or training was rejected. Also, any tests involving value judgements (so-called appreciation tests) were discarded, as requiring a background of musical referents of a specific type for their performance. Typically, such tasks involve the subject in listening to a 'correct' and a 'mutilated' version of an extract of classical music. He must say which version is 'better'. In this situation the ability to discriminate between versions, even to a high degree, does not enable the subject to choose the 'correct' version. Past musical experience is essential to enable him to do this.

The other main omission, compared with existing batteries, is a test of 'harmony' or 'chord analysis' (in which a subject must say how many tones are present in a particular chord). The reasons for this omission are as follows: (*a*) Children below the age of 10 years do not discriminate between consonance and dissonance (Shuter, 1968; Bentley, 1966; Valentine, 1913). (*b*) The normal untrained ear perceives complex sounds and chords differently from the trained ear. Children with formal training are taught analytic listening procedures which enable them to perform certain tasks better than those without such training (Teplov, 1966; Stumpf, 1914). (*c*) Children's perception of how many tones comprise a chord is a function of the tonal composition of the chord itself (Teplov, 1966; Stumpf, 1914; Hickman, 1969). The untrained ear hears more notes in some chords than in others containing the same number. The untrained ear is not necessarily an unmusical ear however.

Nature of the sound source

The accuracy of a person's judgements, particularly in pitch matching (Ritsma, 1967; Ritsma & Engel, 1964) and chord analysis tests (Hickman, 1969), is affected by the wave-form of the stimulus material. Familiarity with a particular kind of

sound (timbre) is also a determining factor (Teplov, 1966). Use of musical instruments (e.g. piano, organ) may introduce differential, and not merely constant, error. In a pitch matching task (in which the possibility of matching by 'beats' was eliminated) Davies (1969) also showed that the accuracy of the judgements of a musically trained group was not significantly affected by change of timbre, whereas a musically untrained group showed significantly poorer performance with certain wave-forms. It is desirable therefore to use tones of standard non variable wave-form throughout; and desirable that the chosen form be produced from some source other than a musical instrument. Musical instruments tend to produce tones of varying timbres at different frequencies, and some test constructors use different timbres on different parts of their test batteries. Sine-wave was chosen, as this does not often occur in chunks of any considerable length in normal orchestral or instrumental music (though some woodwind instruments can produce fairly pure sounds on notes of certain frequencies), and it is unlikely that a musically trained group will ever have performed musical tasks using sine-wave material.

Sound sources consisted of Advance Type J-2, Marconi Audio-Tester, and Ahrend Van Gogh TV-1A type, signal generators, producing sine-wave. Stimulus items were recorded on Revox tape-recorders, using Ampex Professional tape at 7½ in./sec. A tape 'bank' of frequencies was built up, and test items constructed by splicing segments of appropriate length.

The test battery

Melody test

The classic 'melodic memory' test involves listening to paired presentations of a tonal sequence. The second presentation characteristically contains *one* changed note, which subjects are asked to locate (Wing, 1948; Bentley, 1966; Seashore, 1938). The emphasis is thus placed entirely on a single tonal element. To perform the task correctly, subjects must detect a change of *pitch* in one element on the second presentation. The task is thus a pitch discrimination test in which (*a*) the subject does not know beforehand the exact temporal location of the comparison tones, and (*b*) the 'discrimination' tones are surrounded by 'noise'. Factorial studies often show high loadings on a common factor for pitch and melodic memory tasks; however, if a person has difficulty in discriminating between two tones, he will also have difficulty in the melodic memory task. A new type of material was sought, therefore, which would place emphasis on memory for entire sequences rather than individual elements. (The 'courbe mélodique' (Teplov, 1966) remains basically unchanged even when single notes are altered.) When a child sings 'Three Blind Mice' the tune is still 'Three Blind Mice' when a fair proportion of the constituent tones are out of tune or 'wrong', and we can recognize the jingle despite the pitch changes. We certainly cannot assert without qualification that the child does not remember the tune.

The final version of the melody test was as follows. Subjects were first presented with a tonal sequence of three or four tones. After a pause, subjects heard a longer tonal sequence of four, five, six or eight tones. Sometimes the short sequence was contained intact in the longer one. Subjects had to indicate those trials in which the long sequence contained the shorter sequence by answering 'Yes' or 'No'. Also, a location measure was used, whereby subjects indicated the position of the shorter sequence on those trials they judged to be positive.

In the final version, stimulus tones were of 0·63 sec. duration. Initially, tones were presented at a rate of 2 per sec.; this was found to be too fast for many children. Inter-stimulus time was 2·26 sec., and inter-item interval (answer interval) was 3·33 sec. There were 15 items.

The tonal sequences used were statistical approximations derived by an adaptation of Miller & Selfridge's (1953) procedure for producing statistical approximations to English language. Bias in favour of any one musical style was eliminated; certain other test batteries have been biased in favour of classical music (Wing, 1948; Lowery, 1926). Wing, for example,

expressed the opinion that [illegible] is largely confined to [illegible]. However, the [illegible] should be the construction of [illegible] rather than the selection of [illegible] test or profile. With regard to the present battery, items were based on the equal tempered scale. Where attempts would involve the production of sequences based on a scale [illegible] to any culture in an attempt to reduce experiential bias still further.

Pitch test

Pitch tests characteristically involve paired presentations of two tones. Subjects must say whether the two tones have different pitches, and are often asked to indicate the direction of change (does the second note go 'up' or 'down'?) on positive items (Mainwaring, 1931; Seashore, 1938; Kwalwasser & Dykema, 1930; Wing, 1948). Some workers have used smaller than semitone differences (Seashore, 1919, 1938; Bentley, 1966); this may lead to difficulties in noisy rooms, or where electronic equipment does not run at absolutely constant speed. 'Same' items tend to be judged 'different', since the criterion for discriminating between 'same' and 'different' tones is fairly stringent.

The above type of task was excluded from the present battery, since the ability to detect a simple change of pitch does not correlate highly with other 'musically meaningful' tasks (Teplov, 1966). A change of pitch involves changes of frequency, intensity and timbre—all these serve as cues in detecting change of pitch. Laboratory experimental studies suggested that a pitch *matching* task would discriminate well between musical and non-musical groups however.

In an experimental pitch matching task (Davies, 1969) subjects were presented with tones of various kinds; upon cessation of each tone, subjects had to find a tone of the same frequency using an audio oscillator. This kind of procedure was adapted to suit the needs of a group-testing situation in the following manner. The subject first hears a sine-wave stimulus tone; this is followed by a 'glide' or 'sweep frequency' tone, which is divided into 'segments' by short periods of no signal. Subjects must localize the stimulus tone in the segment within which it occurs (i.e. in the segment which traverses a tone of that frequency). In the final version, the glide tone comprised two segments, and stimulus tones could occur in either or neither of the segments.

Stimulus tone duration was 3·2 sec. Glide tone sweep was from 500–1300 Hz, split into two segments by 1·3 sec. of silence, interjected at 880 Hz. Each segment of glide tone was of 2·3 sec. duration. Sweep speed on the final version was governed by an 8 r.p.m. Crouzet clock motor, driving the tuning spindle of a Marconi Audio-Tester. There were 15 items; 3·33 sec. was allowed for answering.

Interval test

An extensive study of interval is given by Madison (1942). The use of interval tests has been less widespread than other types of test material. Also, the great majority of work using intervals has treated them as an element of harmony (two tones struck simultaneously). Typically, subjects have been asked to name intervals, or to discriminate between them according to a 'same-different' paradigm, in a paired-comparison situation. This type of presentation uses interval as an element of harmony; reasons have already been given why any test using harmonic material might be inappropriate in the present testing situation.

In the present study, successive presentation was used. Whilst simultaneous presentation may be viewed as an aspect of harmony, the successive presentation of intervals cannot be viewed as the reduction of melody to its simplest form. Teplov (1966) reports that Meissner showed that children could reproduce the 'courbe mélodique' of a tune heard once, but that they failed in the accurate reproduction of the intervals. Brehmer (1925), Stern (1927) and Gesell & Ilg (1946) have also shown the difference between perception of melodic shape and perception of interval.

The present test also involves a transpositional element. Lundin (1949) used a transpositional task of melody, but did not isolate memory for tonal sequences from ability to transpose tonal sequences. The task devised here presents transposition in its simplest form, by using subjects' ability to compare the *pitch ratio* of two consecutive tones with the *pitch ratio* of two other consecutive tones. In simple terms, the subject compares the difference between the pitches of two tones with the difference between the pitches of two other tones.

In the final version of the present battery, subjects are first presented with two consecutive

tones. After a short pause [illegible] period, two more consecutive tones are heard. The second tone of each pair is higher than the first, and the two first [illegible] tones of each pair are different intervals in musical notation, ranged from one semitone to a [illegible]. The subject compares the tonal separation of the tones in the first pair with the tonal separation of the tones in the second pair, and then indicates which pair has the greater tonal separation, or whether tonal separation is the same in both pairs, giving three possible answers per item. Stimulus tones were of [illegible] sec. duration. Inter-interval time was 2.26 sec., and 3.33 sec. was allowed for answering. There were 15 items.

Rhythm test

Several different approaches to tests of rhythm occur in different test batteries. Seashore (1919) used paired comparison of rhythmic patterns, in which subjects had to say whether versions compared were the same or different. Wing (1948) used comparison of similar paired tunes, with subjects asked to judge the appropriateness of two modes of varying rhythmic accent. More recently, Bentley (1966) used paired comparison of rhythmic sequences, in which the emphasis was on subjects' ability to locate a changed element in the second sequence, in much the same way as elements are located in many tests of melody (melodic memory). A thorough study of rhythmic abilities has been performed by Thackray (1969), in a factor analysis he found that the best single test, out of those he examined, was one in which the subject was asked to reproduce the rhythm of a melody.

The test proposed in the present battery differs in two important ways from any of the above. Firstly, the results found by Thackray are not used here because factor analyses of musical tests suggest that separate factors exist for rhythmic, as distinct from melodic, material (Shuter, 1968). It was decided, therefore, to look for a rhythm test which did not confound rhythm with melody. This attempt to isolate variables was considered important, and follows the example of Bentley, who removed the confounding factor of rhythm from his melodic memory test. This is important, as the degree to which memory for melody facilitates memory for rhythm, and *vice versa*, is not known. Certainly, good memory for tonal sequences is likely to be a help rather than a hindrance in tasks of memory which utilize melodic material. Secondly, the present author is in disagreement with established opinion as to what are the *necessary* characteristics of a rhythmic pattern. Basically, most writers have viewed rhythm as simply what is left of a tune if all tonality is removed, i.e. we are left with elements which differ in intensity (accent), which have variable time intervals between them, and which themselves are of variable duration. Whilst the first two of these sources of variation are essential (it is argued here that they in fact define the rhythm totally), the variation in length (duration) of elements is seen as a characteristic of tones, or noises, and not a characteristic of rhythm. Rhythm is seen instead as an order which the listener imposes upon sequences of tonal elements solely on the basis of their relative intensity, and their relative times of onset. It is argued also that changes in duration of elements *in no way* change rhythm, provided that accent and relative time of onset do not change.

If we hear a symphony played on two occasions at different speeds, we do not say that the rhythm has changed. Note also that if we play a tune on an *instrument* (i.e. all tones have duration) and then ask a person to *clap the rhythm* (in which all elements have the same, short duration) this is a meaningful task. If the person shows that he has correctly perceived the onset times of the different elements, he is judged to have perceived the rhythm correctly.

Finally, if the relative onset times are a crucial feature in the perception of rhythm, the longer we make each element, the 'noisier' becomes the presentation; since, if onset is a 'signal', the period of 'no signal' becomes progressively more obliterated as duration of elements increases, and the harder the task (viewed as an auditory discrimination task) becomes. Thackray himself, from his own results, concludes that tests involving duration are less reliable than those that do not. One can envisage the absurd example of 'abstracting' the 'rhythm' from a tune played in a very 'sostenuto' style and obtaining nothing but a continuous signal (effectively 'no signal'). Rhythm, then, is an order imposed by the listener on the basis of accent and onset times, independent of duration of elements. A fuller exposition of these arguments is given in Davies (1969).

In summary, the view here is that rhythm is a system of temporal anticipations created by the listener, i.e. the belief that certain events will take place at certain *specific* future *points in*

time. If the listener cannot anticipate in this way, for him there is no rhythm. 'Filled' or 'empty' space between points does not fundamentally affect this process.

The nature of the task selected for inclusion in the present battery is as follows. Subjects are asked to compare rhythmic sequences (produced by tapping on a wood block) of varying degrees of complexity with certain regular 'metres' produced on a metronome. In the metres, *the first beat of each measure is accentuated by a synchronized bell*. First, subjects are presented with six measures of a certain metre (duple, triple or quadruple). After a pause, a simple or complex rhythmic pattern is heard, presented at the same tempo as the previous metre. Subjects must say whether the metre and the rhythmic pattern fit together, answering 'Yes' or 'No'. No ambiguous rhythms are included.

Regardless of number of beats in a bar, six measures of metre are always presented. A pause of 3·2 sec. follows, during which subjects write how many beats there are per measure. The comparison rhythmic pattern then follows. There were 15 items; 3·33 sec. was allowed for answering.

RESULTS

In its final form the test battery was administered to 537 children, aged from 7 to 11 years, attending schools in the Durham City area. Sex distribution of the total sample was 46·18 per cent male and 53·82 per cent female. Subjects' responses were marked on standard answer forms involving a minimum of verbal material. Instructions were pre-recorded on the test tape. Duration of the battery was 30 min. and 15 sec.

Table 1 gives means and standard deviations for the whole group, together with possible total scores, for each subtest.

Table 2 gives the mean scores for each age group, for each subtest, together with tertile scores. Expected guessing score for the battery is 21·6. Scores on all parts of the test show small but consistent increases with increasing age.

Table 1. *Means, standard deviations and totals possible for all subjects* (n = 537)

	Melody	Pitch	Intervals	Rhythm	For all tests
Mean	11·73	7·78	7·31	7·16	33·93*
S.D.	3·07	2·68	2·78	3·27	8·52
Total possible	22	15	15	15	67

* Discrepancy due to rounding errors.

Table 2. *Mean scores* (*in bold face*) *and upper and lower tertiles for all age groups*

	Melody	Pitch	Intervals	Rhythm
7+ to 8 years (n = 106)	8·78	5·69	6·08	4·16
	9·76	**6·41**	**6·65**	**5·6**
	11·8	7·88	8·21	7·67
8+ to 9 years (n = 145)	9·92	6·67	6·22	6·07
	10·96	**7·35**	**7·21**	**6·72**
	12·67	8·57	8·52	8·73
9+ to 10 years (n = 139)	10·43	6·91	5·9	6·6
	11·51	**7·805**	**7·33**	**7·49**
	13·25	9·33	8·9	9·7
10+ to 11 years (n = 147)	11·6	7·58	6·86	7·34
	12·77	**8·68**	**7·99**	**8·33**
	14·36	10·22	10·28	10·28

Item analysis. An item analysis, using the double tetrachoric method (Vernon, 1948) with corrections for non-median dichotomization, was performed. This revealed that two items in the melody test and two in the intervals test did not discriminate very well; the remaining 56 items were satisfactory.

Test–retest reliability. Ninety-seven children performed the tests on two separate occasions. Four months and one week elapsed between trials. Children in each age range were tested on both occasions. Results here are given in Tables 3 and 4 for children aged 7+ to 9 years and 9+ to 11 years, respectively.

Table 3. *Test–retest reliability for children aged* 7+ *to* 9 *years* ($n = 60$)

Melody	0·73
Pitch	0·412
Intervals	0·37
Rhythm	0·46
Total battery	0·701

Table 4. *Test–retest reliability for children aged* 9+ *to* 11 *years* ($n = 37$)

Melody	0·76
Pitch	0·58
Intervals	0·54
Rhythm	0·63
Total battery	0·82

Table 5. *Scores on subtests and whole battery for selected criterion groups*

	Melody	Pitch	Intervals	Rhythm	Total
Non-musical adults	13	8·91	6·91	8·57	38·1 ($n = 11$)
Musical adults	20·4	13·75	11·5	13·62	59·46 ($n = 8$)
Non-musical children	12·9	8·71	8·4	9·01	39·02 ($n = 13$)
Musical children	15·65	11·75	12·25	10·75	51·45 ($n = 20$)

In the younger group, reliability is rather poor (0·70); in the older group it compares with reliability coefficients of other workers, 0·82 (Wing, 1948; Bentley, 1966).

Validity. Due to time limits imposed on the study, no longitudinal data are available. Validity data come chiefly from comparison of the performance of criterion and non-criterion groups, and correlation with scores on the Bentley tests. Comparison was made of small groups of musical (choir-boys) children and non-musical (randomly selected) children; and musical (performing) and non-musical (randomly selected) adults. Table 5 gives the scores on subtests and whole battery for each group. t tests of the six possible comparisons of mean total scores were all significant ($P < 0{\cdot}01$), except the comparison between non-musical adults and non-musical children, which was not significant. Despite the small numbers, due mainly to difficulties in obtaining musical criterion groups on a 'performance' basis, results are most encouraging.

Correlations between individual subtests in the present battery and in the Bentley battery (1966) were low, ranging from 0·26 (rhythm) to 0·42 (pitch). This is not surprising when the difference in the nature of the test material is considered. However, a product moment correlation between total scores on the two batteries of 0·66 was obtained. We may thus conclude that if the Bentley tests measure musical ability, then the present battery also does to a considerable degree.

Attempts to obtain reliable teachers' rankings failed. Judgements seemed too often to be obscured by a halo effect from performance in other subjects. However, a rank correlation of 0·68 was obtained between test scores and judgements by a choir master of a group of choir boys.

CONCLUSIONS

Within the limits imposed by the restricted nature of the sample, it is clear that *a battery of tests using essentially non-musical material is feasible.* The test–retest reliability is comparable with that given for other batteries in the older age groups. This falls off somewhat with the youngest children; but no comparable data are available at the moment from other tests within this age group.

No long-term data are available on the validity of the battery, but discrimination between criterion and non-criterion groups appears to be excellent, and there is a significant correlation with the Bentley test. It is also clear that high scores are not dependent upon any formal musical training, since people without training can, and do, obtain high scores.

Correlations between the different subtests are fairly low, ranging from 0·256 to 0·362. A simple factor analysis of the subtests (Thurstone's centroid method) suggests that one factor accounts for most of the common variance. The pitch test yielded the highest loading of 0·645, and the rhythm test the lowest with 0·5. A common factor thus appears to underlie the tests, but the specificity of individual subtests is more striking than their communality. Further studies of the test's performance and its relation to other variables are in progress.

REFERENCES

ANASTASI, A. (1961). *Psychological Testing.* New York: Macmillan.

BENTLEY, A. (1966). *Musical Ability in Children, and its Measurement.* London: Harrap.

BREHMER, F. (1925). In Teplov (1966).

DAVIES, J. B. (1969). An analysis of factors involved in musical ability, and the derivation of tests of musical aptitude. (Ph.D. thesis, Durham University.)

GESELL, A. & ILG, F. (1946). *The Child from Five to Ten.* London: Hamilton.

HEVNER, K. & LANDSBURY, J. (1935). *Oregon Musical Discrimination Tests.* Chicago: Stoelting.

HICKMAN, A. (1969). Musical imaging and concept formation. (Paper read at Conference on Research in Music Education, Reading.)

KWALWASSER, J. & DYKEMA, P. W. (1930). *Kwalwasser-Dykema Tests.* New York: Carl Fischer.

LOWERY, H. (1926). Cadence and phrase tests in music. *Br. J. Psychol.* **17**, 111–118.

LUNDIN, R. W. (1949). The development and validation of a set of musical ability tests. *Psychol. Monogr.* **63**, 305, 1–20.

MADISON, T. H. (1942). Interval discrimination. *Archs Psychol.* 268.

MAINWARING, J. (1913). Tests of musical ability. *Br. J. educ. Psychol.* **1**, 313–321.

MILLER, G. A. & SELFRIDGE, J. A. (1953). Verbal context and the recall of meaningful material. *Am. J. Psychol.* **63**, 176–185.

MURSELL, J. L. (1937). *The Psychology of Music.* New York: Norton.

RITSMA, R. J. (1967). Frequencies dominant in the perception of the pitch of complex sounds. *I.P.O. Annual Progress Report,* no. 2, 5–10.

RITSMA, R. J. & ENGEL, F. L. (1964). Pitch of frequency-modulated signals. *J. acoust. Soc. Am.* **36**, 1637–1644.

SEASHORE, C. E. (1919). *The Psychology of Musical Talent.* New York: Silver Burdett.

SEASHORE, C. E. (1938). *Psychology of Music.* New York: McGraw-Hill.

SHUTER, R. (1968). *The Psychology of Musical Ability.* London: Methuen.

Stern, W. (1927). In Teplov (1966).

Stumpf, C. [illegible] Über neuere Untersuchungen zur Tonlehre. [illegible]

Teplov, B. M. (1966). *Psychologie des aptitudes musicales*. Paris: Presses Universitaires de France.

Thackray, R. (1969). An investigation into rhythmic abilities. *Music Education Research Papers*, no. 4. London: Novello.

Valentine, C. W. (1913). The aesthetic appreciation of musical intervals among school children and adults. *Br. J. Psychol.* 6, 190–216.

Wing, H. D. (1941). Musical ability and appreciation. (Ph.D. thesis, London University Library.)

Wing, H. D. (1948). Tests of musical ability and appreciation. *Br. J. Psychol. Monogr. Suppl.* no. 27.

(*Manuscript received 14 October 1970*)

Br. J. Psychol. (1971), 62, 4, p. 567
Printed in Great Britain

PAPERS TO APPEAR IN FORTHCOMING ISSUES

(Not previously listed)

DIXON, N. F. & HAMMOND, E. J., Department of Psychology, University College London. The attenuation of visual persistence.

EVANS, J. St B. T., Department of Psychology, University College London. Reasoning with negatives.

GOODWIN, R. Q. & WASON, P. C., School of Cultural and Community Studies, University of Sussex, Falmer, Brighton. Degrees of insight.

HAMILTON, P., HOCKEY, G. R. J. & QUINN, J. G., MRC Applied Psychology Research Unit, Cambridge. Information selection arousal and memory.

HAMILTON, V., Department of Psychology, University of Reading. The size constancy problem in schizophrenia: a cognitive skill analysis.

HASLAM, D. R., Department of Psychology, University of Bristol. Field dependence in relation to the influence of stimulus scale-interval upon the assessment of pain threshold.

INNES, J. M., Department of Psychology, University of Birmingham. Word association response commonality and the generation of associative structures.

KIRSNER, K., MRC Developmental Psychology Unit, Gordon Street, London. Developmental and ageing changes in short-term recognition memory.

LAWSON, E. A., School of Studies in Social Sciences, University of Bradford. Vertical disparities.

MAYO, P. R. & BELL, J. M., School of Cultural and Community Studies, University of Sussex, and Education Department, London Borough of Croydon. A note on the taxonomy of Witkin's field-independence measures.

MORRIS, P. E. & REID, R. L., Department of Psychology, University of Exeter. Canadian and British ratings of the imagery values of words.

NICHOLSON, J. N. & GRAY, J. A., Department of Experimental Psychology, University of Oxford. Peak shift, behavioural contrast and stimulus generalization as related to personality and development in children.

O'CONNOR, N. & HERMELIN, B. M., MRC Developmental Psychology Unit, Gordon Street, London. The reordering of three term series problems by blind and sighted children.

REMINGTON, R. E. & STRONGMAN, K. T., Department of Psychology, University of Exeter. Operant facilitation during a pre-reward stimulus: differential effects in human subjects.

ROSS, H. E. & REJMAN, M. H., Department of Psychology, University of Stirling. Adaptation to speed distortions under water.

TAYLOR, J. G., 110 Freedom Avenue, Riverside, Kitwe, Zambia. Underwater distortion: and plain distortion.

VINCE, M. A. & CHINN, S., Psychological Laboratory, University of Cambridge. Effects of external stimulation on the domestic chick's capacity to stand and walk.

Br. J. Psychol. (1971), 62, 4, pp. 549–558
Printed in Great Britain

BOOK REVIEWS

Decision and Stress. By D. E. Broadbent. New York and London: Academic Press. 1971. Pp. xiv + 522. £7.25.

In this new book Broadbent takes as his starting point his earlier work, *Perception and Communication* (1958) and the theoretical model for information processing within the human organism which he at that time proposed. Most experimental psychologists are familiar with its main features: a short term store in which information from the senses is held, a selective filter which only passes some of the incoming information, a limited capacity channel which can only deal with information at a certain rate, followed by a long term store for information processed by the limited capacity system. The new book deals with developments since 1958 in the study of vigilance, selective perception, reaction time, memory and response to stress. The main features of the original theoretical model are retained, albeit with certain important additions, translations and changes of emphasis to accommodate new experimental findings.

This is not a book for beginners. Each chapter is a closely argued review of a large area of experimental psychology from Broadbent's personal viewpoint and each chapter requires some knowledge of the topics before Broadbent's achievement in putting together such an overwhelming mass of evidence can be fully appreciated. The book will undoubtedly serve as a handbook for researchers during the next few years in the same way as *Perception and Communication* served during the last decade.

It is arguable, however, whether the model in general is the most notable achievement of the book, rather than the evaluation of evidence in the individual chapters. As Broadbent readily admits, the reason why the model can accommodate such a variety of experimental evidence is because of its generality. But in its general form it is more of a conceptual aid than a model of how anything within the nervous system actually works. Before considering the model, however, one must assimilate the new terminology.

In the new terminology the filter is retained, acting as a gate on input by allowing information relating to certain features of the stimulus to pass through to the limited capacity channel. Filtering is not now conceived as an all-or-none process – it may involve a hierarchy of tests for stimulus features. After filtering, the information provides 'states of evidence' which are inputs to the limited capacity channel. The limited capacity channel is the 'categorizing' stage and its output is a 'category state'. The set of categories which may be used on any particular occasion are referred to as 'pigeon-holes' and 'pigeon-holing' is the term introduced to describe the manipulation of the set of categories into which the output from the limited capacity channel can fall. Thus filtering is the selection of inputs with certain stimulus features and occurs prior to the limited capacity channel, whereas pigeon-holing determines the possible states of the output from the limited capacity channel. Selection between outputs without discriminating between inputs would be a case of pure pigeon-holing.

A glossary of the terminology is given at the beginning of the book, but although it provides useful illustrative examples it is a mixed blessing if taken too literally. Not all the terms defined in the glossary are used in precisely the same way throughout the text. Even the definition of 'stimulus', 'stimulus feature' and 'stimulus event' can lead one into logical tangles. How many features need to be analysed when the stimulus for one reaction is 'a single square or two circles' and the stimulus for the alternative reaction is 'a single circle or two squares'? The answer in the text (p. 188) seems to be 'two' but one would not suspect this from the glossary definition of 'stimulus feature'. Again, the limited capacity channel is specifically referred to as the categorizing stage (e.g. p. 16) and categorizing is conceived as a process distinct from filtering. If one turns to the glossary for 'categorizing' one finds: 'the process by which the nervous system adjusts so as to allocate certain category states to certain stimulus events'. But 'stimulus events' are not supposed to have direct access to the limited capacity channel: the input to the channel is a 'state of evidence' which arises from a stimulus event but only after filtering (and random disturbance). This confusion, between 'categorizing' used to describe a distinct process only carried out by the limited capacity channel and 'categorizing' used as a general term to include filtering and other information processing, tends to recur – for example, categorizing 'changes the rule connecting a *set of stimuli* to a set of responses' (p. 187).

This does not make the reader's task any easier, but the confusion is not likely to be seriously misleading since it is usually clear from the context what sense is required.

In the discussion of selective perception, an identification is made with Anne Treisman's terminology. Thus her 'dictionary units' correspond to 'category states' and the occurrence of a particular category state corresponds to firing of a dictionary unit. Treisman's 'signals' are Broadbent's 'evidence', varying thresholds for dictionary units is 'pigeon-holing' and Treisman's 'attenuation' is Broadbent's 'filtering'. Broadbent considers that the advantage of his terminology is that it is more general and less likely to be misinterpreted in terms of direct physical analogy. Many readers may incline to the opposite point of view. Terms such as 'attenuation' and 'threshold' have physical interpretations which are capable of experimental test. They imply a specific model which must be changed if the evidence does not fit. It is much more difficult to define what experimental results would constitute a disproof of the applicability of the more general terms. 'Filtering' itself, when it was first introduced, seemed to have a direct physical interpretation, which was subsequently shown to be untenable as an explanation of results from dichotic listening situations. Scientific progress is more likely to be made in this way than by making one's terminology more abstract to circumvent such difficulties.

This is not to quibble about Broadbent's massive achievement in studies of selective attention (as well as in many other fields) but reflects an uneasiness about his determination to apply the terms 'filtering' and 'pigeon-holing' in a highly general way as explanatory concepts. For example, the detailed evaluation of studies of vigilance, the use of signal detection measures and the circumstances which determine changes in either d' or β, would make a substantial monograph on its own. But anyone who has mastered the discussion will have already moved beyond the general concepts of filtering and pigeon-holing. Again, in the discussion of primary memory, the use of semantic classes to distinguish wanted from unwanted items is given as an instance of 'pigeon-holing'. In a different chapter the adjustment of reaction times to stimulus probabilities, in studies such as those of Laming, is also described as 'pigeon-holing'. Will this new terminology help researchers to relate work in these two fields or will it tend to obscure important differences between them?

When so much has been included in this book it might seem ungenerous to complain of omissions. Nevertheless, one notices the lack of any detailed discussion of the use of coding of redundant signals to reduce input. The work of Barlow and others on coding of sensory signals provides one area where some physiological evidence is available on the way in which the nervous system reduces the enormous number of incoming signals without loss of information. It would be interesting to know how this fits into Broadbent's scheme.

These, of course, are impressions from the first reading of a book which will have to be re-read many times before its full value can be properly appreciated. Broadbent's own capacity seems almost unlimited and it is unlikely that anyone else will produce a work on such a scale for some considerable time. Every experimental psychologist will need to read some of it, most will want to read all. May one hope that a cheaper edition will soon be available so that research students can afford to buy their own copy?

R. DAVIS

Structuralism. By JEAN PIAGET. London: Routledge & Kegan Paul. 1971. Pp. vi + 153. £2.00; paper, 50p.

A new book by Piaget is an event, and this one is no exception. Unlike most of the works translated into English it is an essay in the philosophy and methodology of science, and there is relatively little about developmental psychology. However, beneath the surface it is in fact very much about the Piagetian scheme and the role it can play in increasing understanding of ourselves and the world.

The point of departure is the notion of 'structure' which lies at the root of structural modes of analysis in numerous disciplines ranging from mathematics to anthropology. For Piaget, structure is a system of lawful transformations which is in some sense self-regulating. While structural analysis has proved extremely valuable, it encounters difficulties regarding the problem of origins which may be illustrated with reference to linguistics: the dilemma is between 'structureless genesis' which involves 'atomistic association' of the S–R type, and the viewing of structures as given and innate à la Chomsky. Neither is satisfactory, and throughout the book Piaget hints at rather than says explicitly what his answer would be.

Subsequent chapters survey structuralism in practically all the major disciplines, and one marvels at the breadth of Piaget's erudition. It can only be mentioned that he deals with mathematics and logic, physics and biology, and philosophy. Other chapters are of more direct concern to psychologists, above all the one on psychological structures. This contains a penetrating critique of Gestalt psychology and a highly condensed exposition of his own central ideas, with a passing side swipe at critics like Bruner. In the chapter on linguistic structuralism it is fascinating to follow him grappling with Chomsky, recognizing the problem which led the latter to postulate innate features of structure but suggesting an alternative solution with (so far rather slight) empirical support from the work of Sinclair de Zwaart. The section on the social sciences includes Lewin and small group behaviour and hits at those microsociologists who 'continue to adhere to statistical methods whereby relations are, no doubt, quantified, but not in any sense explained'; and Talcott Parsons is praised for going beyond such empiricism. There are also brief excursions into economics and law, and Lévi-Strauss is discussed in detail – more about this later.

Beneath all this there is another theme that Piaget began to explore some time ago, notably in *Biologie et Connaissance*. There he put forward the view that cognitive processes are continuous with the biological adaptation of the organism to the external world. In itself, as he actually states, such a hypothesis may be somewhat trite; but the inferences he draws from it are far-reaching. Thus Piaget argues that the fit between abstract mathematical systems and physical reality, often discovered after the former had been elaborated (and then occasioning surprise), is due to the adaptive functioning of cognitive processes. The next step in the argument is to show that the cognitive structures analysed by Piaget contain within themselves the germ, as it were, of the most advanced structural theories in mathematics, logic and physics. This is the line which is further pursued in *Structuralism*. For instance, a number of French mathematicians writing under the pseudonym 'Bourbaki' have arrived at what they call a 'mathematical architecture' which led Piaget to comment as follows: 'It is remarkable that, psychogenetically, topological structures antedate metric and projective structures, that psychogenesis inverts the historical development of geometry but matches the Bourbaki "genealogy"!' Similarly, Piaget claims that there are physical structures which correspond to our operational structures although they are of course quite independent of human agency; and in one passage he goes so far as to regard this as 'remarkable proof of that pre-established harmony among windowless monads of which Leibniz dreamt'. These and other correspondences could be considered both as cross-validating the basic Piagetian scheme and offering a key to some fundamental epistemological problems.

While the dazzling sweep of Piaget is, for the most part, very persuasive, some doubts remain. Thus, in the case of Lévi-Strauss, Piaget demonstrates with customary brilliance something the present reviewer has attempted to show in a pedestrian manner in the past, namely the fallacy of Lévi-Strauss's notion of 'natural logic'. However, on having accomplished this he protests that he throws no doubt on the specifically structuralist aspects of Lévi-Strauss's analyses. Therefore Piaget obviously does not realize that having removed the cornerstone of the edifice he leaves it, if not in a state of collapse, at least very shaky. One is thereby prompted to wonder how biologists and physicists would react to Piagetian interpretations; one would like to know the standing of the Bourbaki 'genealogy' in the world of mathematics, and it seems that only someone at least as much of a polymath as Piaget is himself would be in a position to make a proper evaluation of his daring bridge-building.

It should be evident even from this rather inadequate review that there is an extraordinary density of ideas in these mere 150-odd pages. Piaget has also been well served by his translator, who has added a number of helpful notes. The book is a *tour de force*, which should be read not only by psychologists but by a wider scientific and philosophical audience. It is to be hoped that it will stimulate an interdisciplinary debate which may well turn out to have profound repercussions.

GUSTAV JAHODA

Visual Pattern Recognition. By P. C. Dodwell. New York: Holt, Rinehart & Winston. 1970. Pp. x+276. £4.70.

The student who wishes to have the traditional facts about visual pattern recognition explained to him will come away from this book somewhat disenchanted. If he is to achieve what the book has to offer he will have to work for it, and not only within these pages but in the references cited.

The details of the earlier conceptions are seldom described, but are instead commented upon as if the author were engaged in an argument with his academic associates. This makes the book valuable to the 'mature' reader, since he does not have to bear with descriptions he knows well, although he may not be entirely convinced at all stages of the commentary. Redundancy in the text is rare, and the language is scholarly and often assumes a certain level of expertise in the reader.

In parts the work reads like a page in the *Psychological Review*; for example, frequent references are made to the literature without a statement of its content. This is fine for well-known works, but not for references which many would regard as specialized if not obscure. Thus to be sure they support the argument, the reader should check. This suggests again that the major section of the book is addressed to active workers in fields related to pattern recognition.

The content ranges from epistemological considerations of perception and the role of psychology, through most useful reviews of certain topics and detailed models for various stages in the pattern recognition process. Several of the review chapters were excellent. I refer, in particular, to those on the neurological and behavioural evidence in regard to the coding of contours by the visual system, and also to binocular vision, which is nicely described as well as comprehensive.

The models deal with attention, perceptual and discrimination learning, adaptation to visual distortion and initial coding. I liked them, but naturally one is disappointed by a lack of supporting evidence for source. When models are unaccompanied by predicted empirical findings they can still be judged by their coverage of existing data and their neurological plausibility. As far as the former criterion is concerned, we are told quite explicitly how the model can deal with existing data; as far as the latter goes, the relationships are not obvious. But the models are probably in a form suitable for testing and one must hope that this will emerge in later publications.

The big problem in pattern recognition is of course the phenomenon of stimulus equivalence. It is a problem because it entails the mapping of what may be an infinite number of conceivable physical patterns on to what may be a finite number of categories. Once one feels one has come to grips with this problem, the 'understanding' fades and one is left in a state of unease. This is how I felt about Dodwell's explanations. They may be as clear as one can approach at present, yet....

D. W. J. CORCORAN

Advances in Psycholinguistics. Edited by G. B. FLORES D'ARCAIS and W. J. M. LEVELT. Amsterdam and London: North-Holland. 1970. Pp. x+454. £8.40; paperback, £5.00.

This is an extremely expensive but beautifully produced volume. It contains a set of papers presented at a very enjoyable conference in North Italy, organized by the University of Padua. Participants were both European and American, and it would be true to say that the speakers addressed each other and their immediate hearers rather than a potential wider audience. The editors do not attempt to explicate the issues further, limiting themselves to brief introductions to sections of papers. The result is a volume which exactly attains the editors' stated aims: 'Rather, this volume will serve the purpose of directing the reader's attention to various new and lively trends in psycholinguistics' (Preface).

The volume does more than produce under one cover articles which either have appeared or are about to appear in the journals. Rather, it juxtaposes often apparently contradictory findings of psychologists all working in particular areas of research. Moreover, the papers as printed contain the results of the informal discussions which were easily the most profitable feature of the conference. As a result highly stimulating solutions are provided. The solutions are enlightening in at least two ways. They may involve an intellectual *tour de force* (as in the case of Bever's resolution of the apparently contradictory findings of the Clarks and Smith and McMahon on temporal relations). Or they may show how specific task variables may preclude an explanation in 'deeper' linguistic or cognitive terms (as in Wason's comments on papers on deductive reasoning by Johnson-Laird and Legrenzi).

Two basic assumptions underlie most of the contributions: that it is justifiable to talk in terms of linguistic and cognitive competence as opposed to performance; and that these two 'competences' are intimately connected. The first of these assumptions is made explicit in the editors' introduction to the first part of the book: 'Linguistic theories are designed for the explanation of

certain facts of human cognition, and linguistic constructs are psychologically real to the degree to which they contribute to the explanation of the empirical phenomena under consideration.' And later, 'The question, then, is whether certain linguistic notions designed for the explanation of a particular domain of facts (i.e. intuitions about grammaticality, paraphrase, etc.) can be used more extensively for the explanation of *other* cognitive phenomena (i.e. the perception of speech sounds, memory for sentences etc.).' The second assumption, that cognitive and linguistic competences are closely connected and might in fact reflect the same underlying processes, leads to some of the most stimulating contributions. The absence of the notion of communicative competence, except in Slobin's contribution, is disappointing.

Provided that the reader can suspend his reservations about the competence assumption, he will enjoy the arrangement of the volume. The two main sections relate to studies where the grammatical features of language figure in the manipulation of the independent variable; and to work involving analysis of the dependent variable in linguistic terms. In the first section the problems of embedded sentences and ambiguous sentences are treated at length. It is noteworthy that several contributors to this section conclude that general perceptual and cognitive skills rather than specific linguistic structures are required to process such difficult sentences. Bever, for example, notes that in embedded sentences a single phrase can perform two separate functions, and he infers a perceptual difficulty in spotting both. Wright suggests that primacy and recency memory effects may confound a linguistic explanation of her results on responses to active and passive sentences. Kurcz stresses that frequency of usage of different sentence types cannot be ignored completely. Garrett complains that it is difficult to determine whether linguistic selection rules or cognitive 'knowledge of the world' affect the interpretation of a sentence. Finally, Slobin summarizes his cross-cultural developmental findings as follows: 'The important advances in language development thus seem to be tied to such variables as increasing ability to perform a number of operations in a short time, increasing short-term memory span, and increasing cognition of the categories and processes of human experience' (p. 184).

It is thus hardly surprising that the second main section, where task and therefore cognitive variables are explicitly manipulated, is more clear-cut in its implications. The aim is to show 'how general cognitive principles determine certain kinds of performance with linguistic material' (pp. 247–8). Kirk and McMahon show how responses to questions testing comprehension of sentences expressing relations in time vary, depending on when those questions are asked. Their results appear to conflict with those of Eve Clark, who suggests that comprehension follows the development of the understanding of temporal relations *per se*. Bever resolves the difficulty by postulating an interaction between organizing the terms in temporal order and in terms of which is the assertion of the sentence and which the presupposition. Another fascinating area concerns the comparative construction; Herbert Clark stresses different levels of linguistic processing, but d'Arcais finds that when grammatical functions are held constant, there is still an effect due to one term of the comparison being the 'focus'. Finally, deductive reasoning; clear papers by Johnson-Laird and Legrenzi show how the particular experimental technique of requiring selection from alternative sources of confirmation or disconfirmation of a statement leads to considerable confusion among intelligent adult subjects. The whole section reveals how dependent we are on the availability or otherwise of a coherent theory of cognitive development, so that the independent variables in those studies reported are selected according to some rationale, and *post hoc* explanations are less frequent.

Other smaller sections concern meaning and aphasia; the different approaches by the several contributors result in little interaction, but in some interesting individual work.

In summary, this volume is a valuable potential source of stimulation for practising psycholinguists and bright advanced students; it is definitely not (nor does it claim to be) a review of, or an introduction to, the area of psycholinguistics.

P. HERRIOTT

Language Development: Form and Function in Emerging Grammars. By LOIS BLOOM. Cambridge, Mass. and London: M.I.T. Press. 1970. Pp. xiv + 270. £4.20.

This book is a most important addition to the literature on the development of language. It reports research done by the author for her doctoral dissertation at Columbia University in 1968. The aims, methods and results of this work are bound to have a major effect on the course of developmental psycholinguistics and no one working in or near this area can afford to ignore

them. The presentation is exceptionally thorough and detailed and the book is in many ways a model of what a research monograph should be. The very full description of the data makes fascinating reading. Of particular value are the tabulations of the single-word utterances and individual lexical items, together with their syntactic contexts in the longer utterances, for each of the three children, who were between 19 and 27 months old when they were observed. This presentation of individual data enables detailed comparisons between the subjects which uncover interesting individual differences. The author also carefully compares her results with those obtained by Brown, McNeill and others.

Previous research on the acquisition of linguistic ability, and in particular of syntactic ability, has almost always viewed the problem in 'formalistic' terms. The ability to speak was equated with the ability to produce grammatically correct sentences. The innovation, and the importance, of this book is that it is the first work which faces up to the fact that the real task for a child is not just to learn to speak grammatically but to learn to speak sensibly. In acquiring language the child acquires communicative competence of which purely syntactic competence is only a part. In order to study the growth of this ability, it is necessary to take account not only of the utterances that the child emits but also the extra-linguistic contexts in which they are produced. To provide this information the author recorded both the children's words and the non-linguistic behaviour and situational context which accompanied them. This non-linguistic information was then used to infer the communicative intentions underlying the utterances.

Such an unashamedly mentalistic approach will clearly cause some concern, not only among neo-behaviourists. We are bound to ask 'Just how good are the investigator's inferences?' and there is no single or certain way to answer this question on the basis of the data that has been presented. However, the thoroughness and care which the study displays should give a strong measure of reassurance. There also seems to be no real alternative to this reliance on intuition. Chomsky has argued that we cannot make progress by pretending that linguistic behaviour is under stimulus control in the same way as the behaviour of a well-trained rat in a discrimination learning experiment. Clearly, there is a relation between what a person says and the situation in which he perceives himself to be but it is seldom a very elementary or direct one. Because of this, we cannot fall back on any simplistic 'objective' methods to try to discover how other people's words relate to the world. We simply have no alternative but to 'use our heads'. The best confirmation we can expect for any results so obtained is the agreement of independent and qualified judges.

Bloom claims that, from the earliest stages, children's utterances express attribution, location and basic semantic functions such as actor, object and verb. I will not try to summarize her results here. A splendidly clear sketch of them has already been given by Roger Brown in his essay, 'The First Sentences of Child and Chimpanzee', which appears in his recent book *Psycholinguistics*. These claims are very interesting and important and, as the author notes in her discussion, they imply the need for a search for the cognitive basis of early linguistic ability, rather than the invocation of syntactic universals. With ideas and problems like these to explore, psycholinguistics can hardly fail to be exciting in years to come.

R. Q. GOODWIN

Toward Unification in Psychology. Edited by JOSEPH R. ROYCE. Toronto: Toronto University Press; London: Oxford University Press. 1970. Pp. 308. £6.00.

The unification or integration of psychology is something to which psychologists often pay lip-service, but relatively few are willing to do much more about it unless they think that their theory or conceptual stance could provide the basis for the unification. A serious attempt to explore the possibilities for unification, by a diverse group of authors with at least no collective axe to grind, is therefore very welcome. This volume attempts to survey the range of approaches to psychology, with the hope that 'with such full knowledge, we might at least be able to *point* to the ways theoretical integration might follow'. The book is the proceedings of a conference which brought together 'physiological and clinical psychologists, psychophysicists and humanists, philosophers of science and animal men, existentialists and perceptionists' in order to air their views about the nature of psychology and to see if they could delineate any common ground. They were prepared for the possibility that they could not, for 'theoretical integration is where you find it, and if it results in fragmentation, then so be it, for our primary concern is to

understand behaviour rather than perpetuate a historical chapter heading for political or historical reasons.'

The book contains a prologue by J. R. Royce, eight position papers, excerpts from the discussions of the papers, and an epilogue by David Krech. The first three papers are general theoretical. In the first, J. R. Royce attempts a taxonomy of approaches to knowledge; he distinguishes rationalistic, metaphorical (intuitive) and empirical epistemological preferences, discusses (with some empirical data) the relationships between them, and suggests the places they might all have in psychology. W. W. Rozeboom issues a long, clear, imposingly mathematical, metatheoretical call for greater rigour in the construction and criticism of theories, and shows what he means with detailed dissections of several varieties of S–R learning theory. Eugene Galanter argues that all specifically psychological explanation should be in terms of psychological variables only: the effect of physical variables, such as stimulus intensity, is of interest only in applications of theory to practical tasks. The next five papers focus more on specific content than the first three did; the authors present their cases for considering their particular areas as central to psychology. S. H. Bartley argues that only those aspects of behaviour which can potentially be assimilated to biological systems should count as part of psychology. L. von Bertalanffy and L. K. Frank each recommend general systems theory as a widely relevant tool for representing the activities of living organisms. R. B. MacLeod argues that extensive phenomenological description is a necessary precursor to adequate psychological theory, and that psychologists have not yet paid enough attention to the descriptive task. Herman Tennessen, in a somewhat unrelated paper, contrasts the ideal of mental health, that of happy, comfortable, purposeful adjustment to a life situation, with the insights of certain existential theories, in which a genuine understanding of man's condition reveals it as a 'wild, banal, grotesque and loathsome carnival in the world's graveyard'. O. H. Mowrer also participated in the conference, but his paper was published earlier in a separate volume.

This book makes interesting reading; the papers are all good, and some of them are very good indeed. On the other hand, it is not altogether clear what conclusions are to be reached concerning the possibility or impossibility of unification. Each of the papers presents a well-articulated position on how to do psychology, and the positions are subjected to frequently searching criticism in the discussions. However, there is very little attempt to see whether or not, and to what extent, the positions are compatible or incompatible, complementary or independent. There is no published discussion on how some of the positions might or might not be able to combine or coalesce into a greater whole, or on how the positions might be tested against each other to determine which is best. This absence of relevant critical discussion is particularly unfortunate in a book of this sort. We already know the claims of phenomenology, general systems theory, biological analysis, etc., for forming the basis of psychology. What we need is informed discussion about the ways of resolving these various claims, or of determining that they cannot be resolved, and it is precisely this kind of discussion which the book fails to provide.

In his epilogue, David Krech states that the conference did nothing to change his pre-conference conviction that the unification of psychology is impossible except on the most trivial level. The conference, in fact, provided very little evidence either for or against Krech's view. Krech's position may well be correct, but in the present volume the case for fragmentation is allowed to win mainly by default.

B. D. MACKENZIE

Comparative Animal Behavior. By Richard A. Maier and Barbara M. Maier. Belmont, Calif.: Brooks/Cole. 1970. Pp. viii + 459. £5.50.

This book begins with a brief consideration of phylogeny, and the authors then work their way through what they describe as a 'comprehensive, integrated comparison of the behavior patterns of animals'. All organisms from amoeba to man fall within their brief, and their exposition involves three sub-goals. In part 1 they deal with such structural and physiological characteristics as sensory systems, nervous and hormonal systems, locomotion and orientation. This is supposed to build the groundwork for consideration, in part 2, of behaviour patterns according to functional criteria such as feeding, defence against predation, reproduction, parental behaviour, social organization and migration and navigation. In part 3 the motivational dynamics of behaviour are discussed through psychogenetics, learning, higher processes, early experience and

[illegible] These three sections are then, to be cumulative, [illegible] in a final chapter on [illegible] of behaviour. [illegible] the text is [illegible] animal behaviour, though the authors are [illegible] useful as a [illegible] for researchers in animal behaviour. The authors have set themselves a Herculean task. To integrate the currently available information on animal behaviour in the way they suggest would be a task beyond any pair of authors, so it becomes here a matter of assessing how far short of their ideal the authors have fallen.

Though the book certainly contains a wealth of facts on animal behaviour, it is no more comparative than the *Guinness Book of Records*. As expected [illegible] the book turns out to be a review of reviews. Of just over two hundred references, almost [illegible] are review articles or books. Furthermore, the reviewing is done in a non-evaluative way. Fact after fact is churned out, not necessarily in an uninteresting way, but without linkage or integrative discussion. Some of the facts are very basic indeed, as for example, 'A large number of invertebrates are sensitive to the force of gravity', or, 'A brainless earthworm burrows much more slowly than a normal one' and 'Fishes that are blinded lose sensitivity to the shade, color and pattern of their surroundings'. Some of the sentences might even have been written tongue in cheek, such as 'The typical centipede (class *Chilopoda*), which has two legs on each segment, is adapted for speed across open land areas', or 'In [illegible], the male utters a sound that resembles a chuckle during copulation; the function of this sound is not clear'. Here and there the book comes to life, intellectually speaking, but only in the form of footnotes where the authors felt that particular matters are controversial or lack evidence. In keeping with their professed aims it might prove more useful to expand the footnotes and abandon the text.

Possibly it is unduly obtuse to suggest that the text is not comparative, and perhaps it is the authors' intention to imply comparisons merely by presenting information in particular sequences or subsections. But in this case what are we to make of the fact that, for example, echinoderms, crustaceans and baleen whales are all filter feeders? It is true that in the later chapter on learning the authors do consider some of the methodological problems involved in making comparisons of animal behaviour, but here again the problems are argued at a general level and only in relation to learning. At no point do the authors consider comparison as a problem applicable in discussions of social, sexual or any other kind of behaviour, though it must be agreed that this is no less than has been done to date by most other psychologists.

On the whole, the information presented seems to be accurate enough, though the Maiers present uncritically and without comment, for example, the claims of Hess that imprinting is a distinct learning type, and Bitterman's claim to have devised a behavioural phylogeny. In the chapter on psychogenetics they manage to propagate a number of myths. It is not, for example, one of the aims of psychogenetics 'to find the gene locus of the behavior pattern', since most behaviour patterns are polygenically determined. Nor is it a problem for psychogenetics that one cannot separate effects of heredity from those of environment. Psychogenetics seeks only to assign aspects of *variance* to genetic or environmental influence. It is no more difficult either to obtain quantitative measures for behaviour than it is to obtain them for morphological characteristics, though of course there may be some problem concerning the meaning of the behavioural measure.

Because it is intellectually undemanding, the book is unlikely to be of much value to undergraduates. On the other hand, since it contains a lot of information and is reasonably well written, it might serve as a text for A level psychology and zoology students.

JOHN WILCOCK

Psychological Experiments with Autistic Children. By B. HERMELIN and N. O'CONNOR. Oxford: Pergamon Press. 1970. Pp. vi + 142. £3.00.

This book reflects the assets and deficits not only of autistic children but also of the experimental method applied to the analysis of deviant behaviour. It reports a series of well-designed and theoretically relevant experiments which provide new insights into the complex disorders from which these children suffer. The authors rely on the models and methods of the laboratory, but show even more than their usual ingenuity in adapting these techniques to the task of assessing children who are often thought to be untestable with the more orthodox psychometric procedures.

Although most of the work has already been reported in some 14 published articles, the book

[illegible] follow the chronological order of the experiments [illegible] presents the material [illegible] perception, and perceptual [illegible] In addition, there is a [illegible] of the [illegible] and previous [illegible] The [illegible] of the findings which could with advantage have been expanded into a [illegible]. In fact the book as a whole is perhaps too concise and could have provided a [illegible] of the experiments and size of the [illegible] in which the subjects were selected [illegible] of the multiple handicaps from which these children suffer and the lack of adequate [illegible] it would have been useful to have more information on this point. We are also not [illegible] whether the same subjects were used in more than one experiment, whether they were [illegible] and whether or not they had evidence of organic pathology. These details [illegible] been included, if only as an appendix. The reader has been spared a vast array of statistical tables, but the absence of much of the raw data makes it difficult for him to check the [illegible] interpretations of their findings. No index is provided.

The findings themselves are of the greatest possible interest and no psychologist or educator [illegible] concerned with cognitive impairments can afford to neglect them or fail to consider their practical implications for the design of remedial programmes. Unfortunately, it is precisely this concern for practical implications which is lacking in the book itself.

The experiments have consistently highlighted certain types of cognitive impairment, some of which appear to be confined to autism, and can therefore be regarded as specific defects, while others are seen to different degrees in other handicapped children and are regarded as developmental disorders. As examples of the latter, the authors describe differences between normal and subnormal children in shape and orientation discrimination, and in simple seriation tasks. Specific deficits which are characteristic of autistic children include a shorter visual inspection time, and abnormal responses to sensory stimulation, mainly in the auditory but also to some extent in the visual channel. More remarkable still are the findings relating to language and language-like functions. Hermelin and O'Connor report some brilliant experiments which demonstrate marked and specific deficits in the appreciation of syntactically organized material. It appears that these children have better immediate rote memories than normal or subnormal 'controls', but they are just as good at recalling unstructured as structured material. More recent experiments by Frith confirm the difficulties which they experience in perceiving order and meaningful structure in input.

This book will repay detailed study. Not only does it contain findings of fundamental importance to students of cognitive development and deviations, but the practising educational and clinical psychologist will find in these pages many ideas for simple experiments which can be applied to the individual child, whether autistic or not.

PETER MITTLER

Developmental Psychology: a Psychobiological Approach. By JOHN NASH. Englewood Cliffs, N.J.: Prentice-Hall. 1970. Pp. xii + 583. $9.50.

Child Psychology. By D. ROGERS. Hemel Hempstead: Prentice-Hall. 1969. Pp. 477. £4.75.

Of the making of textbooks in child and developmental psychology there appears to be no end, at any rate on the other side of the Atlantic, and one's initial reaction to these two volumes, amounting to over 1000 double-columned pages, is inevitably one of doubt and scepticism. In the event, that reaction is unjustified since each in its own way proves to be a useful addition to the literature at different levels of sophistication. Moreover, the Nash volume is significant in providing a new and unusual frame of reference for the study of development that is likely to strike a responsive chord in the minds of many British teachers of the subject, accustomed in the past to a powerful, somewhat unremitting, American environmentalist tradition.

Nash reflects on the changes in his views on development that have taken place in recent years, mainly involving a modification of the central position of learning in the development of behaviour. His approach is cultivated, well informed and many-sided. Essentially it adds up to the position that while environmental influences are crucial to development, man's biological nature has a good deal to do with the kind of environment he creates, and his capacities for learning are biological characteristics, subject to the limitations of his physiology. The book therefore attempts

to deal in considerable detail with the way in which human evolution and biological equipment direct and organize ontogenetic development, and with the interaction of environment with these biological elements. Others have, of course, adumbrated the idea of interaction as a central concern, but not infrequently prefatory and valedictory obeisances to the concept have proved to be unfulfilled in the main body of the text. The task is difficult, but here is a genuine attempt to bring together material from biology with that of the social sciences. The theme of interaction is adequately sustained, and a good deal of material familiar to students of developmental psychology is presented in an unfamiliar and often illuminating context. Nash explicitly distinguishes developmental psychology from child psychology (the latter being the concern of Dorothy Rogers's book), and in consequence is as concerned with phylogenetic and comparative as with ontogenetic material. He regards man as in a stage of dynamic evolution, emphasizes biological sex differences (there is a remarkably comprehensive chapter on this, though the theme pervades the book), and introduces much material from adult studies on the grounds that development has to be seen as having a goal in adult behaviour, and that its study requires integration with the general body of psychological theory. There is a good deal about values in the book, particularly on the values implicit in scientific investigation, and there is also a courageous apology for the use of speculation occasionally as a respectable scientific occupation. He underlines the need for methodological rigour, though the book does not deal much with methodology, but maintains that where significant human issues (for example, the institutional care of children) are concerned an obsession with methodological purity may prove sterile.

The plan of the book, comprising 20 chapters, is consonant with the writer's general, though qualified, view that biological determinants are paramount in earlier life and social influences more important later. It begins with sections on Efficiency in Development, Individuality, Evolutionary and Genetic Influences, Innate Shapers of Development, Constitutional Psychology, Critical Periods, Imprinting, Sex Behaviour, and after dealing with the areas of perception, emotion, motivation and cognitive processes ends with discussions of Identification, Social Influences on Personality and the Concept of Self. His final Synthesis and Projection expounds a philosophy of development owing much to Maslow's theory of self-actualization, and the last word is given to Bertrand Russell, whom Nash regards as an example of the fusion of the Apollonian with the Dionysian approaches to life.

The book is clearly presented, the material wide ranging, and serious students of the subject will find abundant references that they can follow up. Some may find it rather idiosyncratic, but it has the merit of a serious attempt to synthesize the field within a coherent frame of reference which gives biological factors fuller treatment and higher status than do most other works of this kind.

Dorothy Rogers's book is clearly aimed at a different readership from that of Nash, and would prove a helpful introduction for beginning college students, teachers, parent groups, and W.E.A. classes. She has tried to provide 'a palatable presentation without sacrificing intellectual depth'. Significant names and concepts appear in heavy type to indicate their appearance in a comprehensive glossary, and the fact that 'infanticide' is the first of these in the book is an indication not of the author's views on child rearing, but possibly of the intellectual matrix expected of the student for whom the book is designed. The material is concerned with child psychology in the sense of an attempt to describe the characteristics of the child at various stages of development, and there is a legitimate interest in the practical application of this kind of knowledge in the fields of child care, family life, education, and society generally. The author states her reputable and unexceptionable philosophy of child development in the preface. Concepts and theories are carefully and attractively presented, lavishly and usefully illustrated with case material and quotations from sources, and each chapter ends with suggested questions and activities of varying degrees of relevance. As with most American texts of this kind there are pictures and diagrams, some of which add little except to vary the mode of stimulus and information. However, of its kind this is a lively, readable and reliable book whose interpretative approach is not too culture-bound. Incidentally, the denotation of 'child' is taken very seriously, since one has to search hard for any reference to adolescents.

Both books are too expensive for students in this country, and one hopes that the Nash in particular may appear in paperback form before too long.

E. M. EPPEL

Achievement-related Motives in Children. Edited by CHARLES P. SMITH. London: Basic Books, 1971; New York: Russell Sage Foundation, 1969. Pp. vii + 263. £4·10.

In October 1967 a research conference on 'The Development of Achievement-related Motives and Self-Esteem in Children' was held at the City University of New York. It brought about a confrontation of the three major styles of research on these topics, the approaches of Michigan, the Fels Research Institute, and Yale. One outcome is this book, which consists mainly of four longish research reports.

Virginia Crandall begins her paper on 'Sex Differences in Expectancy of Intellectual and Academic Reinforcement' by discussing how expectancy is to be defined, and she distinguishes three usages of the concept. She goes on to report studies showing that girls throughout the range from elementary school to college have lower expectations of accomplishment in intellectual and academic activities than boys. Objective grounds for such a difference are apparently lacking, and she considers, with all the excitement of a 'whodunit', a variety of possible explanations for the observed sex difference.

Joseph Veroff proposes a three-stage development of achievement motivation made up of autonomous achievement motivation (governed by internal standards of excellence), then a social comparison form (in which achievement is defined by the standards of the child's social group), and finally a phase in which the two are integrated. As measures of these two kinds of achievement motivation he has devised batteries of ingenious tests, derived from Atkinson's theory of achievement motivation, and the bulk of the chapter reports his findings. These support his developmental theory, reveal sex differences in types of achievement motivation, and elucidate children's reactions to success and failure.

Charles Smith's chapter on the origin and expression of achievement-related motives is teeming with empirical findings bearing on McClelland's theory of how achievement motivation develops and on Sarason's theory of the development of test anxiety. Using data collected from 10–11-year-old boys and their parents, he examines the relationship between the two motives and brings in such socialization variables as parental child-rearing values as well as personality variables (self-concept, defensiveness) and performance variables (e.g. self-directed learning). The total picture is extremely complex, and it raises interesting questions regarding the multidimensionality of achievement motivation and the relationship between achievement training and independence training.

Sheila Feld and Judith Lewis, who concentrate on achievement anxiety, report a massive operation aimed at re-modelling the Test Anxiety Scale for Children (TASC). They conduct a painstaking and sophisticated analysis of the data collected from 8875 second-grade boys and girls in 111 American schools and offer a penetrating discussion of the results. The crucial issues they consider are whether TASC scores are distorted by an acquiescence response bias; what the scale measures, as indicated by a series of factor analyses; and the usefulness of a re-scoring of it in terms of its factor structure. It is a methodological *tour de force*, as the editor remarks, with interesting substantive findings thrown in.

There follow three chapters in which John Atkinson comments on the first two papers, Seymour Sarason on the third and fourth, and Howard Moss on all four. In these commentaries the voices of expert experience throw out a challenging flow of ideas on the multitude of theoretical and methodological issues arising. Even though they add to the overall complexity by offering salutary reminders of pitfalls in interpretation, raising awkward terminological questions, and suggesting refinements in measurement and conceptualization, they provide invaluable perspective for a stock-taking both of the particular researches reported and of activity in general in this field. This service is continued in the concluding chapter, in which Smith draws together the major findings of the four studies and relates them to previous research. He also considers their implications for child-rearing and for educational practice, including ability-grouping, desegregation and the use of programmed instruction.

Overall, this book creates a picture of a lively field of research in which methodological advances are being made and there is promise of theoretical integration. Within the childhood age range much enlightenment might come in future from longitudinal and cross-cultural studies, and the book itself acknowledges the importance of linking achievement studies of children with those of adults, and of relating both to what is known of cognitive and personality development in general. It should do much to add to the vitality of work on achievement motivation and anxiety.

ALFRED FLOOK

Temperament and Behavior Disorders in Children. By ALEXANDER THOMAS, STELLA CHESS and HERBERT G. BIRCH. London: University of London Press 1969 Pp. vii + 309. £3.15.

This book is the second based upon the New York longitudinal study of 136 children. To my mind this is an extremely important study. It deals with the behaviour problems that developed in 42 of the larger sample during the first 10 years of their lives. The problems included aggressiveness, tantrums, disturbed sleep, non-participation in school activities, stealing, lying and poor peer relations. The authors observe that over 15 years or more of experience with children they came to be impressed with the contribution made to behavioural development by reactive characteristics of the child, particularly his patterns of temperamental organization. They were increasingly convinced that environmental influences could not accommodate the range and variability in the course of development exhibited by individual children. Nor, they felt, could such factors alone explain the marked differences in children's responses to similar patterns of parental care.

The present study was designed to test under controlled conditions their clinical impressions. Nine characteristics of temperament – activity, rhythmicity, adaptability, approach/withdrawal, threshold of responsiveness, intensity of reactions, mood, distractability and persistence – were studied in relation to the children's evolving behaviour patterns. Both before and after they developed symptoms, groups of the children with behavioural disturbances differed in temperament from those who did not develop such disturbances. The clinical cases, as a group, were characterized by an excessive frequency of either high or low activity, irregularity, withdrawal responses to novel stimuli, non-adaptability, high intensity, persistence, and distractibility. No single temperamental trait acted alone in influencing the course of the child's development. Rather, combinations of traits forming patterns and clusters tended to result in an increased risk for developing behavioural disorders. Differences in types of behaviour disorders and of symptoms, too, were found to be associated with differences in temperament.

The authors show that a given pattern of temperament did not, as such, result in a behavioural disturbance. Deviant as well as normal development was the result of the interaction between the child with given characteristics of temperament and significant features of his intrafamilial and extrafamilial environment.

When the children in the longitudinal study population were still under two years of age, and far in advance of the appearance of any behavioural disturbances among them, one subgroup came to particular attention. These children were variously characterized by their mothers, the interviewers, and all other members of the research team in terms of a series of pejorative labels, ranging from the expression 'difficult children' to 'mother killers'.

When the behaviours of these children were analysed, they were found to be characteristically deviant from the overall sample of children with respect to certain attributes of temperament. These attributes included irregularity in biological functions, a predominance of negative (withdrawal) responses to new stimuli, slowness in adapting to changes in environment, a high frequency of expression of negative mood, and a predominance of intense reactions.

It is of great interest that the organization of temperament in this early identified group of so-called difficult children corresponded closely to the temperamental cluster found from later data to characterize a disproportionately high number of the children who developed behavioural disturbances with active symptoms.

Two other temperamental 'types' were found: children who were characterized as 'easy' and children who were characteristically 'slow to warm up'. Case histories and follow-up information about the three types of child are given in great detail. The book contains much useful information about the parents' reactions to these different types of temperament and their handling of their children.

It can best be left to the authors to state the relevance of this concept of temperament to general clinical theory. It lies, as they put it, 'neither in its sole pertinence for behavior disorders, nor in its displacement of other conceptualizations, but in the fact that it must be incorporated into any general theory of normal and aberrant behavioral development if the theory is to be complete. Existing theories emphasize motives and drive states, tactics of adaptation, environmental patterns of influence, and primary organic determinants. The central requirement that a concept of temperament makes of such generalizations is that they come increasingly to focus on the

dividual and on his uniqueness. In other words, it requires that we recognize that the same motive, the same adaptive tactic, or the same structure of objective environment will have different functional meaning in accordance with the temperamental style of the given child. Moreover, in such an individualization of the study of functional mechanisms in behavior, temperament must be considered as an independent determining variable in itself, and not as an *ad hoc* modifier used to fill in the gaps left unexplained by other mechanisms.'

MARTIN HERBERT

Behavior Disorders in School-Aged Children. By HARVEY F. CLARIZIO and GEORGE F. McCOY. London: Intertext. 1970. Pp. xiv + 519. £4.30.

This is a much-needed text. There is a dearth of books (eclectic in their presentation of material) which deal with behavioural disorders in both the psychological and educational contexts.

The authors state that the text was prepared with the following readers in mind: advanced undergraduates and beginning graduate students who are preparing to become clinical psychologists, school psychologists, school counsellors, special education teachers, and, most importantly, ordinary classroom teachers. The text is eminently readable; and it is packed with useful empirical information (notably summarized in tables reproduced from their original sources) and clinical/intuitive accounts of deviant behaviour.

The authors have organized the chapters into three major sections. The first, 'Developmental and Diagnostic Considerations', deals with the problematical concept of 'normality' from a developmental point of view. The two chapters in this section tackle such issues as: normal behaviour problems, common growth problems, adjustment mechanisms, the stability of deviant behaviour, criteria of normality, the concept of diagnosis, classification, and problems in evaluating children's behaviour. The second section, entitled 'Behaviour Disorders in Children', contains discussions of six 'types' of disorders. These include a description of syndromes, aetiological theories and treatment programmes. There is also an excellent discussion of learning disabilities. The final part concerns itself with therapeutic and preventive measures. There is a reasonably comprehensive coverage of psychotherapeutic and behaviour modification techniques, and in addition an interesting chapter on classroom discipline.

Each chapter is graced by a list of suggested readings and an impressively wide and up-to-date bibliography.

MARTIN HERBERT

Perspectives in Personal Construct Theory. Edited by D. BANNISTER. London: Academic Press. 1970. Pp. xii + 273. £4.00.

The late George Kelly is surely one of the most interesting and irreverent figures in modern psychology. He is a man whose work resolutely refuses to be neatly parcelled up and put away in one of those 'boxes' so beloved by psychologists. As Kelly observed in one of the two papers by himself included in this book of essays: 'Personal construct theory has also been categorized by responsible scholars as an emotional theory, a learning theory, a psychoanalytic theory (Freudian, Adlerian and Jungian – all three), a typically American theory, a Marxist theory, a humanistic theory, a logical positivistic theory, a Zen Buddhistic theory, a Thomistic theory, a behaviouristic theory, an Apollonian theory, a pragmatistic theory, a reflective theory, and no theory at all. It has also been classified as nonsense, which indeed, by its own admission, it will likely some day turn out to be. In each case there were some convincing arguments offered for the categorization, but I have forgotten what most of them were. I fear that no one of these categorizations will be of much help to the reader in understanding personal construct theory, but perhaps having a whole lap full of them all at once will suggest what might be done with them.'

It is fortunate, given all this confusion, that an excellent book such as the present one has come into being. Don Bannister, a lucid and always entertaining interpreter of Kelly's work, and himself a researcher into personal construct theory, invited 10 people to write an essay around the theme of personal construct theory. As he says, the brief was no more explicit than that. For Bannister the book was an experiment to test the hypothesis that personal construct theory is

fertile, and in addition that it has cohering properties. As I construe the essays, written by J. B. Adams-Webber, Don Bannister, Fay Fransella, Dennis Hinkle, Ray Holland, Graham Lemon, Don Oliver, Phillida Salmon, John Shotter and J. M. M. Mair, which have titles ranging from the formidably esoteric such as 'Actual Structure and Potential Chaos: Relational Aspects of Progressive Variations within a Personal Construct System' to the inviting but unpretentious 'Psychologists are Human Too', both hypotheses are confirmed. But then I belong to Bannister's second sort. The book itself is by way of being a simple sorting test like the Kelly Repertory Grid itself. As Bannister says, 'The book may also prove a test of the reader's stance towards modern psychology. If you feel that current psychological undertakings, or a fair proportion of them, hold promise and assiduous work along proven lines will advance the discipline, then the book may well seem to you to be intellectual shadow-boxing. If you have grave doubts as to whether psychologists know what they are about and cannot scan a current journal in psychology without a sinking heart, then the book may seem to you properly preoccupied with basic issues.'

MARTIN HERBERT

Understanding Abnormal Behaviour. By L. E. Cole. Scranton, Pa.: Chandler Publishing Co. 1970. Pp. 736. £5.00.

Any reviewer must be puzzled precisely for whom the message in this fascinating and unconventional book is intended. Apart from the title *Understanding Abnormal Behaviour* the clues in the preface read, 'In recent decades awareness of the light thrown on the human condition by abnormal forms of behaviour makes the study imperative for everyone whose ambition is to deal with other human beings, whether an administrator, therapist or theorist.' The clues as to the particular form of understanding of abnormal behaviour intended by 'the study' we presume lies in the author's admission: (*a*) of having set himself, 10 years previously, the task of translating Pierre Janet's *De l'angoisse à l'extase*, and (*b*) 'in his discovery of how great a gap existed in the English language accounts of abnormality' and his 'ambition to present some of Janet's views to American students'.

It may not be to his discredit that, though he does succeed in communicating much of Janet's particular brand of psychiatry, most of the book, in its eclecticism, is largely Cole. Was it Isaiah Berlin who once dichotomized novelists into those who presented a wide tapestry of life and those who distilled the essence of it in one central theme? If so, then like Tolstoy, Cole belongs to the former group, for his tapestry is immense. The modern student of psychopathology now expects his text to be terse and with a fairly clear-cut theoretical framework. This work is more discursive and sceptical in style, so that numerous excerpts from case studies and summaries of empirical data lie cheek-by-jowl with the author's subjective evaluations, usually couched in the richest imagery – but not always posed in their most economic form.

The book is organized into 12 chapters, the titles of which would lead one to suppose an intention to create an introductory text. There are chapters each on the concepts of abnormality, genetic mechanisms, social factors (including some epidemiology in the U.S.A.), child-rearing, neurasthenia, hysteria, psychosomatic disorders, breakdowns in self-regulation (phobias, obsessional states, etc.), brain lesions and cortical malfunction, the affective psychoses and the schizophrenias. Among fairly numerous references to Charcot, Henri Ey, L. Kolb, A. Noyes, P. Bernard, C. Brisset and S. Kraines, Janet, as one would suppose, figures prominently, particularly in the chapters on neurasthenia, hysteria and breakdowns in self-regulation. The core of Janet's theme is represented as a lowering in integrative force ('psychological tension') of the system of conscious regulation resulting in dissociation of function, splitting off of subsystems of habits, etc. In the case of neurasthenia, the reduction in 'tension' results in faulty strategies of attention, and a focus on priority issues which succumbs to matters hierarchically of less importance. Hence therapy consists of helping the patient to cut his losses, and reassert a perspective over his priorities. Therapeutic intervention to restore integrative control in hysterics is less hopeful. Here a constitutional lack of 'tension' only results in removal of control in one function as a consequence of regaining control in another, etc.

Cole's objective is presumably to develop a theory of the structure and dynamics of the self-system. (A second aim is probably to 'knock' psychoanalysis.) If it were not for the untidy theoretical idiom into which these ideas are cast, this sort of 'systems analysis/cost analysis'

approach by which the therapist helps the patient reassess his priorities, concentrate on his [illegible] reapportion his life's investment portfolio, would [illegible] in the [illegible] reader. But what enables a systems-analyst to work effectively is [illegible] and clear-cut analyses of component functions and their cost-benefits. Janet's model, even when [illegible] in terms of modern control system theory, has not yet the [illegible] of being [illegible] the [illegible] of professional skill. In the end the themes in the book, attractively full of big and small wisdoms, still only leave us groping in the right direction. And since this is a direction towards which the modern student, with his training in cybernetics, would lean in any case, it is far from clear that a return to the vague formulations of it by Janet would be all that profitable.

Where, then, would one place the consumer of this book? One could not advocate it as an introductory text, replacing say Maher or Buss, either for psychology or D.P.M. students. But I can see one important role for it. The pressure on time imposed by the modern psychopathology curriculum often militates against an adequate historical treatment of the subject. This seems to leave the student defenceless against the host of implicit assumptions in aetiological theories. These remain as a residue in our psychiatry rather like the bizarre vestigial rudiments on a traditional army uniform. It seems to me that this text, through its historical analysis, may help the more sophisticated student distinguish between the now defunct fripperies, and help him restitch the remaining component garments into a more serviceable battle dress suitable for present-day theoretical psychiatric requirements. So I would say—good reading for professionals and the more sophisticated student; not an introductory text. A. BURSILL.

Psychopathology: the Science of Understanding Deviance. By JAMES D. PAGE. Chicago: Aldine. 1970. Pp. 482. $9.95.

According to Brendan Maher (prefacing his own textbook, *Principles of Psychopathology* in 1966) an understanding of the phenomena of disordered behaviour must always depend on an understanding of the principles of ordered behaviour; yet in the presentation of abnormal psychology to university students (he says) it is surprising how seldom these principles and methods are reflected. The result is that the student is often left confused and dismayed. Maher's approach is still a model in its notice of contemporary experimental methods and results. In 1971 psychopathology (or 'abnormal psychology', as many prefer) might well expect to hold an even firmer place in behavioural science. But there remains the difficulty of its presentation to undergraduates. So many of its concepts are traditional and complex—that some psychologists doubt the wisdom of even attempting to present them at an introductory level. In spite of a slow stream of textbooks in recent years (mostly from the United States) it is hard to name the ideal accompaniment to a fundamental course in the subject. Buss's *Psychopathology* comes nearest perhaps.

Now Page offers *Psychopathology: the Science of Understanding Deviance*, described on the jacket as 'a basic textbook for first courses in psychopathology, abnormal psychology or the psychology of deviance'. Does this suggest an anxiety about 'keeping in' with all shades of psychologicalopin ion? No great matter, perhaps, but unfortunately this feeling of uncertainty of direction persists well into the tome itself, which is a substantial one, in size, anyway. Now, according to Buss (in the introduction to *his* textbook), for this sort of material 'exposition should be rigorous and systematic, observations should be separated from explanation, data from influences and facts from theories'. For all Page's emphasis in his opening sentence that psychopathology is a behavioural science, this is something not always apparent in his book. His approach is avowedly eclectic, which results in his falling between all manner of stools. For instance, though fashionably firm about rejecting the concept of mental disease, he shows uncertain appreciation of the arguments for doing so: *vide* his latter solemn listings of diagnostic categories like 'asthenic personality', 'neurasthenic neuroses', etc. from the American Psychiatric Association classification. Indeed his closest allegiance throughout seems to be with psychiatry; and though he would not have us accept psychiatric classification 'uncritically and *in toto*', he nowhere gives real coverage to any other possibilities. For example, he sketches in Eysenck's theory of neurosis, but never quotes Eysenck's work, or any other work, as an alternative basis for studying behavioural abnormality. This one-sidedness is emphasized by the way he patterns his material to accord with his own preconceived plan. Having taken at the outset the thesis that all behaviour ranges over a continuum, from normal to psychotic, with the personality disorders and neuroses (in this order) in between, he seems to need to confirm and

underline this thesis at every turn. While he is entitled to his preferred approach, he nowhere examines critically alternative interpretations, based on the effects of different treatments, experimental findings, etc. I think a thoughtful student would notice this and be dissatisfied. On the other hand, a less thoughtful one might become dangerously self-assured in the light of some of the rather dogmatic statements made on the basis of unevaluated research to which full references are not invariably given. For British students difficulties could also arise from the differences between American and British psychiatric terminology. But does even American psychiatry equate 'lycanthropy' with 'depersonalization'? And would not the phrase 'acted out' in a definition of 'psychodrama' in the glossary be liable to confuse even the American student struggling with the term 'acting-out' (in the text, but not in the glossary)? Indeed, the glossary seems an altogether arbitrary, unevenly useful list. Many words seem purely medical, lacking any special psychological connotation, e.g. 'atrophy', defined as 'wasting away or shrinking of a bodily organ'.

To conclude, I found this a rather shapeless, loosely pointed book. There is much factual information here but it does not seem as well digested as it might be, and in the way it is presented, its relevance for students of psychology is obscured. I could not recommend it as an accompaniment to any British undergraduate course.

S. WHITAKER

Fundamentals of Behaviour Pathology. By R. M. SUINN. New York: Wiley. 1970. Pp. 508. £4.40.

This is a tersely phrased, highly readable book for undergraduate students, bringing together several approaches into a fresh, invigorating, sensible construction. Each chapter has informative photographs and tables, a glossary of terms and a bibliography. It is difficult to assess the book's value to the British market; the inclusion of references to elementary topics would imply that it was designed to stand as a self-contained course. But in Britain it is still the more usual practice for the course in psychopathology to be given once the student has reached a fair degree of psychological sophistication. Nevertheless, the confluence of elementary principles together with psychopathological themes may achieve some useful consolidation even amongst our more senior students.

The brief historical introduction through the vicissitudes of myth and magic titillated even my jaded palate. The embarrassment of dealing with the student's primitive prejudices regarding fact and attitude are dealt with in chapter 2 by a device little employed recently – the Platonic dialogue; Mr Ignoramus is answered by wise Socrates in a kind of therapeutic psychodrama. The theme is then developed into definitions of abnormality and a useful introduction to the epidemiology of mental ill-health in the U.S.A.

Part 2 contains the main pedagogical innovation of the book. Teachers of abnormal psychology who have not yet found a teaching bridge between general psychology, experimental psychopathology, and descriptive psychiatry could do worse than to emulate this paradigm, for the model is as successful in the lecture room as it is in print. The understanding of personality structure and dynamics is attempted through five chapters, successively on the family, sociocultural, biophysical, psychological and genetic factors. The introductory paragraphs are admittedly sufficiently elementary to make one blush (the introduction to the biophysiological chapter elicits a veritable vasodilatory storm!). Yet, in an odd way, the material does not diminish the author, but rather succeeds in setting the scene. Perhaps the design of this section has not yet fully reached its target, but promises a potentially successful formula for teaching the student various fundamental approaches to aetiology.

Part 3 – 59 per cent of the book – deals with the major adult disorders divided into six chapters, together with a chapter each on child psychiatry, mental deficiency, the therapies and the relation between personality theory and psychopathology. I found these uneven and a little conventional, failing to carry through expectations from part 2. The chapter on schizophrenic disorders is perhaps the best. Descriptions of the conditions are usually followed by a consideration of aetiological theories and the most effective forms of treatment. The terse style is maintained, but I fear some of the chapters are too superficial and eclectic to match the appetite of a good honours student. Nevertheless, teachers of psychopathology should give careful consideration to many instructive innovations in the book. It may also find an 'ecological niche' in some postgraduate social science courses.

A. BURSILL

An Introduction to Social Psychiatry. By RANSOM J. ARTHUR. Harmondsworth: Penguin Books. 1971. Pp. 168. 35p.

Very gradually, there is accumulating a body of empirical fact and tested hypothesis which would justify the separation of social psychiatry as a distinct subspecialty, with its roots as deep in the sciences of psychology and sociology as in the sciences more traditionally regarded as basic to medicine. No text so far available even begins to do justice to this emerging discipline. Dr Arthur's Penguin does make a beginning. He deals with the right issues: the nature of mental illness, the contribution of epidemiology, the significance of social class, the importance of social precipitating factors, the issues arising from transcultural comparisons and the substitution of community-orientated for institution-based services. He writes well. He dislikes vague or grandiose formulations. Although he is not as critical as this reviewer would like, he presents a balanced argument by virtue of his selection of authors, and he does not restrict himself to the American literature. At 35 new pence, his book is a bargain.

J. K. WING

Psychology in the Industrial Environment. By M. P. FELDMAN. London: Butterworth Group. 1971. Pp. viii + 128. £2.20.

Until recently, British psychologists have tended to lag behind their American counterparts in their facility for condensing a wide range of complex material and presenting it in a readable form for non-specialists. In this book on the applications of psychology in the industrial environment, Dr Feldman demonstrates that the skill is not confined to the other side of the Atlantic. An informative guide to such diverse topics as personnel selection, the measurement of morale, design of equipment, the working environment, personality and accidents, and advertising and the consumer, is presented in 11 chapters, none of which exceed 14 pages in length. Inevitably the scope and level of the book preclude the treatment of any topic in depth and, to this extent, the claim on the fly-leaf that 'the important problem of fitting the man to the job is considered in depth' is misleading. The subject is given wider coverage than other topics in that separate chapters are devoted to the questions of aptitudes for management, for selling and for scientific creativity respectively, but they represent an overview of the relevant factors rather than a detailed discussion.

According to the foreword, the book is intended as an introductory text, particularly for students training for, or working in, applied fields for which some knowledge of occupational psychology is desirable but who are not necessarily psychologists themselves. For those with access to a library where they can follow-up the references at the end of each chapter, many of which are to articles in a variety of learned journals, or who will use the text as an adjunct to lectures and tutorials given by an experienced occupational psychologist, the book can be commended. It is slim, light, has a direct style, a good index, and a pleasing format which should facilitate its constant use, although the price will deter many students from purchasing a copy for themselves. For those hoping to use the book unaided, however, one must be less sanguine. The sheer compression and necessary selection of material entailed in its composition has resulted in, at best, an over-simplified and, at worst, a one-sided picture of several issues which will give the impression that our knowledge is more clear-cut than it really is. For example, in a short section on shift work and related personal and social problems, it is stated that where it is necessary for individuals to work rotating shifts 'the period between changes of shift should be long enough to permit adaptation and workers should have several days off between changes'. This is one point of view, but no mention is made that others hold that there is sufficient research evidence to show that the alternative of rapidly rotating shifts may often be preferable on both personal and physiological grounds. It is desirable that businessmen and managers should be aware of this, just as it is desirable that medical students and lawyers should realize that not all factories are organized on the assembly line system, although they will not learn of it here. If the book will stimulate readers to discover more about the extensive and complex topics to which it refers, well and good; but one fears that too many will regard it as a definitive statement of current knowledge of the subject. For its size it seems grossly over-priced, bearing in mind the increasing number of substantial good paperbacks on the market, and one can but regret that in trying to cater for a wide readership too much has been attempted in too small a volume.

SYLVIA SHIMMIN

Radical Man: the Process of Psychosocial Development. By CHARLES HAMPDEN-TURNER. Cambridge, Mass.: Schenkman. 1970. Pp. xii + 434. £2.55.

This unusual book, which took four and a half years to write, embodies the first theoretical dissertation ever submitted at the Harvard Business School. The major part of it consists of a selective review of research findings on dogmatism, conformity, obedience, anomie, the effects of T-groups and other topics which the author considers relevant to an elucidation of contemporary social and political processes and events, notably mass protests against the Vietnam war and student activism. These findings are discussed within the framework of a 'model of psychosocial development' which is, in fact, simply a set of quasi-metaphysical propositions about the nature of man, using the terminology but not the concepts of existentialism, Marxism and social psychology. In the main, the research findings are presented lucidly and accurately, and the inferences drawn from them, though frequently laced with rousing quotations from poetry, novels and newspaper clippings, are seldom iconoclastic. In Dr Hampden-Turner's vocabulary, 'radical' is closer in meaning to 'self-actualizing' than to 'Marxist', and his own ideological viewpoint, which he makes no attempt to conceal, is surprisingly establishmentarian.

The inferences which are drawn from the invariably slender data frequently involve breathtaking extrapolations from the psychological laboratory to the barricades which will stimulate lay readers, but which many psychologists will find tiresome. At times these conceptual long-jumps seem hardly justified by the data: for example, when the author draws the 'tentative conclusions' from Berkowitz's famous experiments on the effects of justified and unjustified filmed violence, that 'the New Left probably *decreases* the amount of violence in the world in the longer run, whenever it succeeds in placing violence in the context of a sickening injustice and "over-reaction"' (p. 279).

The book begins with a full-frontal attack on traditional social science, or what the author calls 'the borrowed toolbox of Conservative Man', but the arguments are trite. This is the weakest section and in any event it does not seem really to belong with the rest of the book. Dr Hampden-Turner argues against the hypothetico-deductive method, which he considers inherently 'manipulative', and in favour of induction. This chapter is in part an essay in the philosophy of science, but no attempt is made to confront the devastating arguments against induction put forward by David Hume, Karl Popper, or any other important critic. More seriously, he seems at times fundamentally to misunderstand what he is supposed to be attacking, as when he asserts that neo-behaviourists 'believe that "reinforcement" under certain "stimulus conditions" will "shape the drives" of specified types of people so that their "responses" can be predicted' (p. 20).

On the whole, this book makes light and interesting reading. The author's modest achievement lies in the quantity of information he has assembled and organized, not the 'model of psychosocial development' which serves simply as the arbitrary framework of organization. But not many psychologists will agree with the publisher's blurb: 'Hampden-Turner most convincingly restores human conscience and moral impulse to the very centre of social science, from which they were banished years ago by behaviourism'. Those who are seeking such a restoration will be disappointed.

ANDREW M. COLMAN

Social Relationships. By G. J. MCCALL, M. M. MCCALL, N. K. DENZIN, G. D. SUTTLES and S. B. KURTH. Chicago: Aldine. 1970. Pp. ix + 193. $7.95.

This is an attractive, if somewhat light-weight, examination of social relationships. The emphasis is on the interaction between dyads rather than groups. The approach is mainly theoretical. Little empirical evidence is quoted to support the suggestions – often intuitively very sensible made by George McCall and his co-authors.

After a consideration of the dynamics of personal relationships by George McCall, their boundary rules are analysed by Michal McCall in an approach which owes much to Goffman. Norman Denzin writes on how conventional rules of behaviour are relaxed in enduring social relationships and the topic of friendship, which is perhaps under-emphasized by psychologists, is examined in an illuminating way by Gerald Suttles and Suzanne Kurth.

Social Relationships may prove to be a helpful introduction to an area which has until now been of more concern to sociologists than to psychologists.

ANGELA B. STEER

Attitude and Attitude Change. By HARRY C. TRIANDIS. New York and London: Wiley. 1971. Pp. xiii + 232. £3.50; paper, £1.65.

This book was written to serve as a multi-purpose, multi-level text. Its intended uses range from relatively intensive coverage of one area in an introductory social psychology course to an introduction to social attitudes for beginning graduate students. It also endeavours to cover the topic of attitude and attitude change quite comprehensively.

These objectives are achieved to different degrees. The book cannot be faulted on the breadth of its coverage, although the depth with which some problems are treated is another matter (to which further reference will be made). In this context, the main virtue of Triandis's book as a general text on social attitudes is the absence of serious competition at the moment. There are texts on attitude theory, others on attitude measurement, contributory volumes based on original research, and books of readings in the attitude area generally; however, the few general texts tend to be either outdated (Cohen), or more elementary and less comprehensive (Zimbardo and Ebbeson). The multi-purpose, multi-level aspirations are less clearly achieved. The very comprehensiveness of the book makes questionable its interest and value for the introductory-level student of social psychology. The expense of even the cheaper paperback edition does not help, as an instructor would have to have his students buy similar books on other course topics, and possibly books of primary-source readings besides. (This is not so much a criticism of the reviewed volume as of the concept behind the series to which it belongs.) The book would be more suitable for a more specialized and advanced undergraduate course, or at an introductory graduate level where it could be supplemented by extensive reading from primary sources and the integrative efforts of the instructor. Stylistically, it is well enough written to hold the already interested reader, but one would hardly use it to seduce the uncommitted.

Triandis deals with some topics exceptionally well. Among the strong points of the book are his chapter on attitude measurement, and his detailed analysis of the relationships among attitudes, norms, behavioural intentions, and behaviour. On the other hand, the one page devoted to such artifacts as demand characteristics in attitude experiments is inadequate, and unlikely to satisfy the critical student. (In fairness to Triandis, he does at least refer such students to other sources.) Coverage in the area of authoritarianism and related sociopolitical attitudes could also be improved upon. The use of study questions and projects for advanced students is laudable in principle, although the author's specific assignments might sometimes be questioned. His study questions tend towards the regurgitative.

The author's reasoning is occasionally difficult to follow. For example, in his discussion of how the affective component of attitudes is formed, he shifts abruptly and without explanation from Schachter's work on the cognitive and physiological determinants of emotion to Scott's research on how rewards affect attitude change in a role-playing situation. Elsewhere, having defined aggression as 'behavior whose goal is the injury of some person or object', he treats social distance as an aggressive response (which it sometimes is, but not surely always).

Occasional errors can also be found. For example, Triandis represents Rosenberg as predicting and finding subjects to express more 'botherment' at generally inconsistent attitudinal cognitions than at personally inconsistent ones. In fact, Rosenberg predicted and found the opposite.

The book's greatest weakness is its lack of theoretical integration. Theory and relevant research are presented in abundance, but critique, comparison and reconciliation are in short supply. While it could be argued that the study questions and projects should act as a Socratic goad to propel the student himself in that direction, most of the assignments are too specifically focused to encourage such global thinking.

Despite the weaknesses referred to, the strengths of this book – its comprehensiveness and excellent treatment of certain key topics – weigh heavily. Its main advantage, however, remains the relative absence of competition in its field.

PAUL KOHN

Freud and Psychology. Edited by S. G. M. LEE and MARTIN HERBERT. Harmondsworth: Penguin Books. 1970. Pp. 398. 50p.

Group Processes. Edited by PETER B. SMITH. Harmondsworth: Penguin Books. 1970. Pp. 454. 60p.

Creativity. Edited by P. E. VERNON. Harmondsworth: Penguin Books. 1970. Pp. 400. 50p.

Thought and Personality. Edited by PETER B. WARR. Harmondsworth: Penguin Books. 1970. Pp. 447. 60p.

The role of an anthologist is an important one, especially when source material becomes scarce as student numbers increase. Yet the anthologist's task is obviously difficult. He has to discriminate between the available studies in his field, and identify (or even create) landmarks. There are probably some similarities between making collections of papers, or extracts from books, and collecting other things, like paintings. To read straight through an anthology is as tiring as it is to go slowly through a big museum, since one satiates rapidly. But anthologies should not be read straight through, as they have an important function in giving structure to a field, and illustrating it with specific examples. In this they are of particular value to beginning students, to whom they can give a grasp on original reports. Although anthologies entail a rigorous selection from the available material, the basis of selection is not always well justified, and 'somebody else' would be certain to select differently. These anthologies show the expertise of their editors, but only Vernon has clearly listed the principles on which he made his selection, and only his introduction is not closely aimed to students at an elementary level.

On the present evidence, the pattern for the Penguin Readings is now quite stable: there is a modal six-page introduction, a mean of 23·5 papers classified to six sections (or nine in the case of *Thought and Personality*), and brief lists of 'further reading' with some useful abstracts in *Freud and Psychology*. It is the ready availability of the papers in these sets that mainly recommends them. At around 2p for each paper, with the editorial material free, they are all worth recommending for relevant courses. The titles give a reasonable guide to what each collection is about, and a casual skim might help an initially disinterested person decide if he could arouse enthusiasm for any of the fields.

Freud and Psychology and *Creativity* are the most coherent and give broadly useful summaries. The other two appear more as 'collections' of papers, with somewhat less unity. Lee and Herbert have used extracts (or excerpts) less than have any of the others. But the ideal strategies are hard to specify, and detailed criticism of the actual selections is inappropriate, especially if a basic aim to convey the character of a field is accepted.

L. B. BROWN

Semantic Differential Technique: a Source Book. Edited by JAMES C. SNIDER and CHARLES E. OSGOOD. Chicago: Aldine. 1969. Pp. xiii + 681. $12.50.

This is described as 'the first source book of readings on the semantic differential technique', a tool that has become one of the most popular in psychology. It has an autobiographical introduction by Osgood, in which he expresses surprise that his technique should have become so successful. It is good to know this is his opinion. The papers that one needs to refer to are here, but none that have criticized the method: there is also a 40-page bibliography and a semantic atlas for 550 concepts.

L. B. BROWN

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review)

AVEDON, E. M. & SUTTON-SMITH, B. *The Study of Games*. Chichester: Wiley. 1971. Pp. xiv + 530. £4.40.

BAIRD, J. C. *Psychophysical Analysis of Visual Space*. Oxford: Pergamon Press. 1970. Pp. viii + 321. £4.50.

BARCLAY, A. M., CRANO, W. D., THORNTON, C. & WERNER, A. *How to do a University*. Chichester: Wiley. 1971. Pp. ix + 191. £1.75.

BARKER, P. *Basic Child Psychiatry*. London: Staples Press. 1971. Pp. 236. £1.50.

BARNES, C. D. & KIRCHER, C. *Readings in Neurophysiology*. Chichester: Wiley. 1968. Pp. viii + 482. £4.20.

BARON, R. A. & LIEBERT, R. M. *Human Social Behavior: a Contemporary View of Experimental Research*. Homewood, Ill.: Dorsey Press. 1971. Pp. xv + 559. $6.95.

BATTEGAY, R. *Psychoanalytische Neurosenlehre: eine Einführung*. Bern: Huber. 1971. Pp. 214. 21 D.M.

BECKER, P. E. *Humangenetik: ein kurzes Handbuch in fünf Bänden: Band 1/1*. Stuttgart: Thieme. 1968. Pp. xv + 534. 198 D.M. + Band 1/2. 1969. Pp. 572. 224 D.M.

BLACKMAN, S. & GOLDSTEIN, K. M. *An Introduction to Data Management in the Behavioural and Social Sciences*. Chichester: Wiley. 1971. Pp. vi + 104. £3.75; paper, £2.75.

BLOS, P. *The Young Adolescent: Clinical Studies*. London: Collier-Macmillan. 1970. Pp. xxiii + 252. £4.50.

BOTT, E. *Family and Social Network: Roles, Norms, External Relationships in Ordinary Urban Families*. (Second edition.) London: Tavistock. 1971. Pp. xxx + 363. £3.25; paper, £1.50.

BOTWINICK, J. *Cognitive Processes in Maturity and Old Age*. New York: Springer. 1967. Pp. x + 212. $5.75.

BRITT, S. H. *Consumer Behavior in Theory and in Action*. Chichester: Wiley. 1970. Pp. xxv + 499. £5.25.

BRITT, S. H. (ed.). *Psychological Experiments in Consumer Behavior*. Chichester: Wiley. 1970. Pp. xv + 416. £3.10.

BUMPASS, L. L. & WESTOFF, C. F. *The Later Years of Childbearing*. Princeton, N. J.: Princeton University Press. 1970. Pp. xv + 167. £3.60.

BURGESS, R. L. & BUSHELL, D. *Behavioral Sociology: the Experimental Analysis of Social Processes*. London: Columbia University Press. 1969. Pp. xiii + 418. £5.65.

BUTTER, C. M. *Neuropsychology: the Study of Brain and Behavior*. Belmont, Calif.: Brooks-Cole. 1968. Pp. xi + 211. Price not given.

CHOMBART DE LAUWE, M.-J. *Un Monde Autre: l'Enfance: de ses représentations à son mythe*. Paris: Payot. 1971. Pp. 443. 46 FF.

COFFEY, H. S. & WIENER, L. L. *Group Treatment of Autistic Children*. Englewood Cliffs, N.J.: Prentice-Hall. 1967. Pp. xii + 132. £2.00.

COTGROVE, S., DUNHAM, J. & VAMPLEW, C. *The Nylon Spinners: a Case Study in Productivity Bargaining and Job Enlargement*. London: Allen & Unwin. 1971. Pp. 151. £2.75.

DICKS, H. V. *Fifty Years of the Tavistock Clinic*. London: Routledge & Kegan Paul. 1970. Pp. xiv + 415. £4.20.

ELLIS, G. (ed.) *Units, Symbols and Abbreviations: a Guide for Biological and Medical Editors and Authors*. London: Royal Society of Medicine. 1971. Pp. 36. 54 np.

ERIKSON, E. H. *Identity: Youth and Crisis.* London: Faber & Faber. 1971. (First published 1968.) Pp. 336. £1.20.

FINNEY, D. J. *Probit Analysis.* (Third edition.) London: Cambridge University Press. 1971. Pp. 333. $18.50.

FOUCAULT, M. *Madness and Civilization: a History of Insanity in the Age of Reason.* London: Tavistock. 1971. (First published in Great Britain 1967.) Pp. xiv + 299. £1.00.

FRIEDMAN, L. J. *Virgin Wives: a Study of Unconsummated Marriages.* London: Tavistock. 1962. Pp. 161. £1.25.

FROMM, E. *Man for Himself: an Enquiry into the Psychology of Ethics.* London: Routledge & Kegan Paul. 1971. (First published 1949.) Pp. xiv + 254. 90 np.

FURTH, H. G. *Piaget and Knowledge: Theoretical Foundations.* Hemel Hempstead: Prentice-Hall. 1969. Pp. xvii + 270. £4.00.

GREENBAUM, S. *Verb-Intensifier Collocations in English: an Experimental Approach.* Hague: Mouton. 1970. Pp. 97. 14 gldrs.

GREENO, J. G. *Elementary Theoretical Psychology.* London: Addison-Wesley. 1968. Pp. 324. £1.50.

HALLER GILMER, B. VON. *Industrial and Organizational Psychology.* (Formerly called *Industrial Psychology.*) 1971. (Also 1961, 1966.) Pp. xviii + 682. £5.75.

HAYES, J. R. (ed.). *Cognition and the Development of Language.* Chichester: Wiley. 1970. Pp. x + 370. £5.25.

HERRIOT, P. *Language and Teaching: a Psychological View.* London: Methuen. 1971. Pp. vii + 109. £1.25; paper, 60 np.

HÖPP, G. *Evolution der Sprache und Vernunft.* Berlin: Springer. 1970. Pp. 167. 38 D.M.

HOWE, M. J. A. *Introduction to Human Memory: a Psychological Approach.* London: Harper & Row. 1970. Pp. ix + 113. £1.40.

HOWELL, W. C. & GOLDSTEIN, I. L. *Engineering Psychology: Current Perspectives in Research.* New York: Appleton-Century-Crofts. 1971. Pp. xiv + 648. $16.50.

HUNT, W. A. (ed.). *Learning Mechanisms in Smoking.* Chicago: Aldine. 1970. Pp. 237. $8.95.

JACOBS, J. *Adolescent Suicide.* Chichester: Wiley. 1971. Pp. xi + 147. £4.50.

JOHNSON, D. A. *Comprehension and Clear Thinking.* Chichester: Wiley. 1971. Pp. xi + 122. £1.00.

JOHNSON, P. E. *Psychology of School Learning.* Chichester: Wiley. 1971. Pp. vii + 124. £2.50.

JOYCE, C. R. B. (ed.). *Psychopharmacology: Dimensions and Perspectives.* London: Tavistock. 1971. (First published 1968.) Pp. ix + 430. £1.50.

JUNG, E. & VON FRANZ, M.-L. *The Grail Legend.* London: Hodder & Stoughton. 1971. (First published 1960 in German.) Pp. 452. £3.15.

KAHAN, V. L. *Mental Illness in Childhood: a Study of Residential Treatment.* London: Tavistock. 1971. Pp. 218. £3.00.

KARLINS, M. *Psychology and Society: Readings for General Psychology.* Chichester: Wiley. 1971. Pp. xi + 514. £5.00; paper, £2.65.

KEEN, E. *Three Faces of Being: Toward an Existential Clinical Psychology.* New York: Appleton-Century-Crofts. 1970. Pp. xii + 367. $4.25.

KING, R. A. (ed.). *Readings for an Introduction to Psychology.* (Third edition.) Maidenhead: McGraw-Hill. 1971. Pp. 584. £2.40.

KINTSCH, W. *Learning, Memory and Conceptual Processes.* Chichester: Wiley. 1970. Pp. xii + 498. £4.40.

KLEIN, M., HEIMANN, P. & MONEY-KYRLE, R. (eds.). *New Directions in Psychoanalysis: the Significance of Infant Conflict in the Pattern of Adult Behaviour.* London: Tavistock. 1971. (First published 1955.) Pp. xiii + 534. £1.05.

KÖHLER, W. *The Task of Gestalt Psychology.* Princeton, N.J.: Princeton University Press. 1969. Pp. vii + 164. £3.15.

LAING, R. D. & COOPER, D. G. *Reason and Violence: a Decade of Sartre's Philosophy: 1950–1960.* London: Tavistock. 1971. Pp. 184. £1.90; paper, £1.00.

LANNOY, R. *The Speaking Tree: a Study of Indian Culture and Society.* London: Oxford University Press. 1971. Pp. xxvi + 466. £5.50.

LYMAN, S. M. & SCOTT, M. B. *A Sociology of the Absurd.* New York: Appleton-Century-Crofts. 1970. Pp. xi + 221. $3.30.

MACGAFFEY, W. *Custom and Government in the Lower Congo.* London: University of California Press. 1970. Pp. xv + 322. £4.75.

MAHRER, A. R. (ed.). *New Approaches to Personality Classification.* London: Columbia University Press. 1970. Pp. 425. £5.65.

MANGAN, G. L. & BAINBRIDGE, L. D. *Behaviour Therapy.* London: University of Queensland Press. 1969. Pp. 148. £1.40.

MERBAUM, M. & STRICKER, G. (eds.). *Search for Human Understanding: a Reader in Psychology.* New York: Holt, Rinehart & Winston. 1971. Pp. xi + 531. No price given.

MILLER, N. E. *Selected Papers.* Chicago: Aldine. 1971. Pp. xiii + 874. $15.00; paper, $6.75.

MORGAN, C. T. & KING, R. A. *Introduction to Psychology.* (Fourth edition.) Maidenhead: McGraw-Hill. 1971. Pp. 784. £5.25.

MUNDLE, C. W. K. *Perception: Facts and Theories.* London: Oxford University Press. 1971. Pp. 192. £1.25; paper, 55 np.

NUNNALLY, J. C. *Introduction to Psychological Measurement.* Maidenhead: McGraw-Hill. 1970. Pp. xv + 572. £4.80.

ORD, I. G. *Mental Tests for Pre-literates, resulting mainly from New Guinea Studies.* London: Ginn. 1970. Pp. xiii + 270. £2.75.

PAGÈS, R. & POULAT, E. *L'Orthodoxie religieuse: essai de logique psychosociale.* Paris: Editions Ouvriers. 1971. Pp. 373. 80 FF.

PEPITONE, A. *Attraction and Hostility: an Experimental Analysis of Interpersonal and Self Evaluation.* London: Tavistock. 1971. (First published 1966.) Pp. 238. 75 np.

PINNEY, R. *Creative Listening.* Loddington, Northants.: Creative Listening. 1970. Pp. 8. 25 np.

PITTENGER, O. E. & GOODING, C. T. *Learning Theories in Educational Practice: an Integration of Psychological Theory and Educational Philosophy.* Chichester: Wiley. 1971. Pp. 222. £3.50.

POWELL, E. H. *The Design of Discord: Studies of Anomie.* New York: Oxford University Press. 1970. Pp. xi + 247. £3.45.

READ, K. H. *The Nursery School: a Human Relationships Laboratory.* (Fifth edition.) London: Saunders. 1971. (First published 1950.) Pp. xi + 401. £3.00.

ROGOW, A. A. *The Psychiatrists: the First Inside Look into the Fascinating World of Psychiatry and Psychoanalysis.* London: Allen & Unwin. 1971. Pp. 315. £3.50.

ROSENFELS, P. *Homosexuality: the Psychology of the Creative Process.* New York: Libra. 1971. Pp. 169. $5.95.

ROSENTHAL, D. & KETY, S. S. (eds.). *The Transmission of Schizophrenia.* Oxford: Pergamon Press. 1969. Pp. 433. £6.30.

ROTHMAN, D. J. *The Discovery of the Asylum: Social Order and Disorder in the New Republic.* Boston, Mass.: Little, Brown. 1971. Pp. xx + 376. $12.50.

RUSSELL, W. R. *The Traumatic Amnesias.* London: Oxford University Press. 1971. Pp. ix + 84. £2.00.

RYLE, A. *Neurosis in the Ordinary Family: a Psychiatric Survey.* London: Tavistock. 1971. (First published 1967.) Pp. 156. 75 np.

SAVAGE, R. D. *Psychometric Assessment of the Individual Child.* Harmondsworth: Penguin. 1968. Pp. 117. 30 np.

SCHMIDT, U. *Figurale Nachwirkungen auf kinästhetischem Gebiet bei Lageveränderung von Objekt und Tastender Hand.* Frankfurt am Main: Waldemar Kramer. 1970. Pp. 153. 24 D.M.

SENN, M. J. E. & HARTFORD, C. (eds.). *The Firstborn: Experiences of Eight American Families.* Cambridge, Mass.: Harvard University Press. 1969. Pp. ix + 533. $11.00.

SHABERMAN, R. *Autistic Children.* London: Stuart & Watkins. 1971. Pp. 94. £1.50.

SHANDS, H. C. *Semiotic Approaches to Psychiatry.* Hague: Mouton. 1970. Pp. 412. 58 gldrs.

SIMKINS, L. D. *The Basis of Psychology as a Behavioral Science.* Waltham, Mass.: Ginn-Blaisdell. 1969. Pp. ix + 198. $3.25.

SMITH, R. W. (ed.). *Guilt: Man and Society.* New York: Anchor Books. 1971. Pp. 314. $1.95.

STITES, R. S. & STITES, M. E. *The Sublimations of Da Vinci + Codex Trivulzianus.* Newton Abbot: David & Charles. 1970. Pp. x + 422. £6.30.

STONE, I. *The Passions of the Mind: a Novel of Sigmund Freud.* London: Cassell. 1971. Pp. 808. £2.60.

STORMES, J. M. & CRUMPLER, J. P. *TV Communications Systems for Business and Industry.* Chichester: Wiley. 1970. Pp. xiii + 238. £5.75.

TASHNER, M. *The Failure Hypothesis.* New York: (author), 299 Madison Ave, N.Y. 10017. 1970. Pp. 42. Price: inquiries to author.

THOMPSON, J. D. & VAN HOUTEN, D. R. *The Behavioral Sciences: an Interpretation.* London: Addison-Wesley. 1970. Pp. 268. Price not given.

THOMPSON, W. R. & DEBOLD, R. C. *Psychology: a Systematic Introduction.* Maidenhead: McGraw-Hill. 1971. Pp. xiii + 453. £4.55.

TOCH, H. *The Social Psychology of Social Movements.* London: Methuen. 1971. (First published 1966.) Pp. xiv + 257. 90 np.

TOMAN, W. *Family Constellation.* (Second edition.) New York: Springer. 1969. Pp. viii + 280. $5.75.

VEBLEN, T. *The Theory of the Leisure Class.* London: Unwin. 1970. (First published in U.K. 1925.) Pp. xx + 261. 75 np.

VOEKS, V. *On Becoming an Educated Person: the University and College.* (Third edition.) London: Saunders. 1970. Pp. xxi + 278. £1.25.

WALSH, M. N. (ed.). *War and the Human Race.* Amsterdam: Elsevier. 1971. Pp. 95. Price not given.

WILLS, W. D. *A Place Like Home: a Hostel for Disturbed Adolescents.* London: Allen & Unwin. 1970. Pp. 136. £1.75.

WINNICOTT, D. W. *Playing and Reality.* London: Tavistock. 1971. Pp. xiii + 169. £2.10.

WRENCH, D. F. *Readings in Psychology: Foundations and Applications.* London: McGraw-Hill. 1971. Pp. xii + 456. Price not given.

EDITORIAL ACKNOWLEDGEMENT

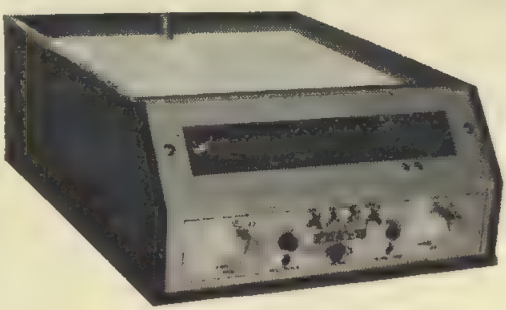
The Editor, Assistant Editor and Co-Editors wish to acknowledge the help they have received in the past year from the following persons, who have acted as referees for papers submitted for publication in the *British Journal of Psychology*.

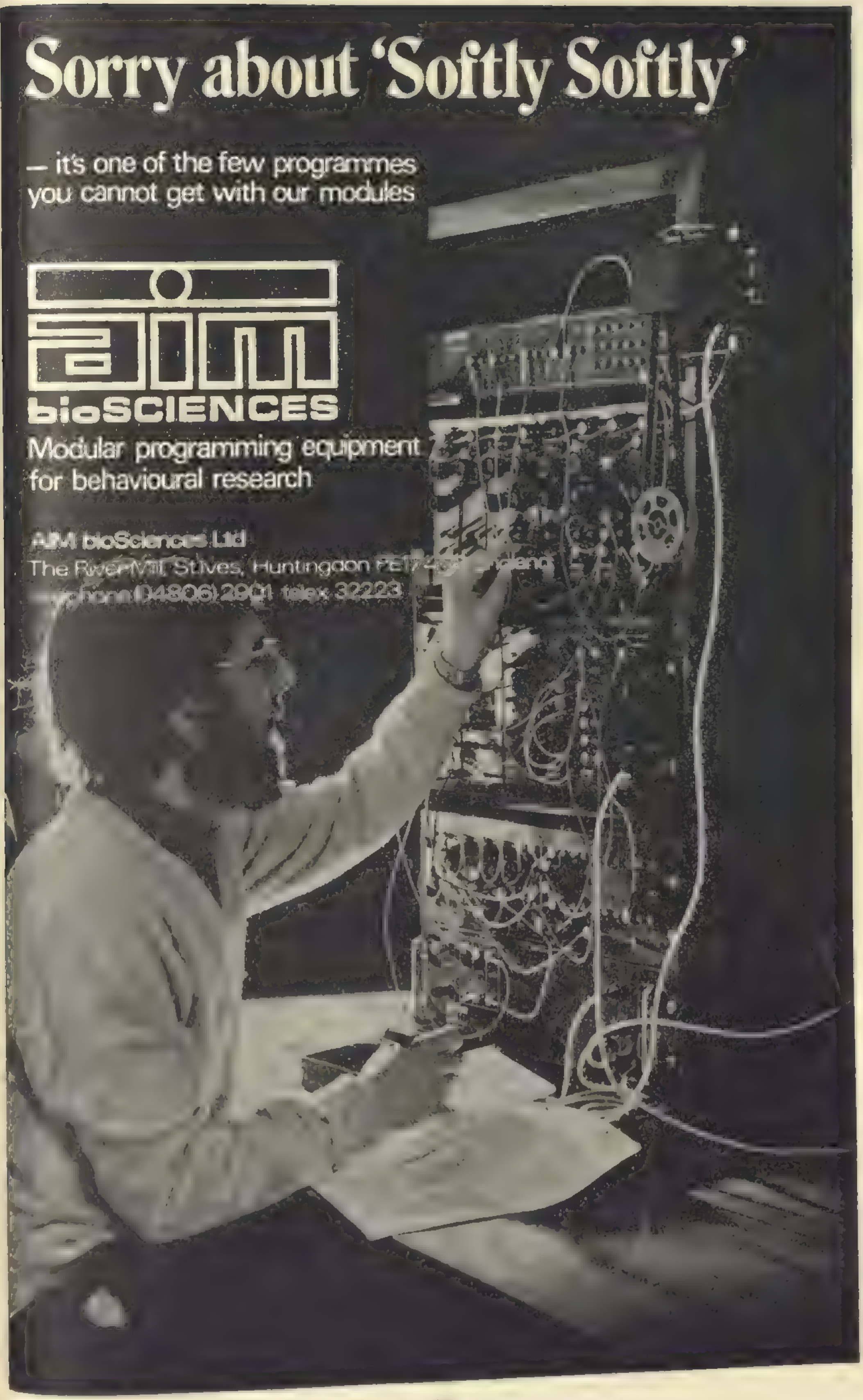
DR S. M. ANSTIS
MRS H. S. ANTHONY
DR A. BADDELEY
DR D. BANNISTER
DR S. A. BARNETT
MR J. H. W. BARRETT
DR HALLA BELOFF
MR R. B. BETT
DR D. E. BLACKMAN
DR D. BOYLE
DR P. G. BRITTON
PROF. P. L. BROADHURST
DR I. D. BROWN
DR R. BROWN
DR W. P. BROWN
DR VIDA CARVER
DR W. CHEYNE
PROF. A. D. B. CLARKE
DR A. M. CLARKE
DR P. COLQUHOUN
DR K. J. CONNOLLY
DR F. I. M. CRAIK
MR W. G. CUMBERBATCH
MRS ANN TAYLOR DAVIES
DR D. R. DAVIES
DR J. D. DAVIS
PROF. G. DREW
PROF. H. J. EYSENCK
DR S. B. G. EYSENCK
DR B. FELLOWS
PROF. D. W. FORREST
DR J. P. FRISBY
DR U. FRITH
MISS C. E. O. FULLERTON
MR A. GALE
DR R. GARSIDE
DR H. B. GIBSON
DR I. E. GORDON
MR M. J. GRAY
PROF. R. L. GREGORY
MR D. GRIFFITHS
PROF. H. GWYNNE JONES
DR M. S. HALLIDAY
PROF. D. W. HARDING
DR G. HARRISON
DR B. HERMELIN
PROF. I. M. L. HUNTER
MR J. HUTT
PROF. G. JAHODA
DR P. N. JOHNSON-LAIRD
DR S. JONES
DR R. B. JOYNSON
PROF. H. KAY
DR C. KIERNAN
DR G. R. KISS
DR P. KLINE
MR C. R. LEACH
DR. D. LEGGE
DR J. A. LEONARD
MR A. B. LEVEY
DR P. M. LEVY
DR S. A. LEWIS
MR B. MANDELL
MR W. A. MATTHEWS
DR W. S. MATTHEWS
PROF. J. MCLAUGHLIN
DR P. MITTLER
DR J. MORTON
DR B. MOULDEN
PROF. J. D. NISBET
DR N. O'CONNOR
MR W. ORWIN
DR I. OSWALD
DR J. P. PHILLIPS
DR MARY PICKERSGILL
MR G. W. PILKINGTON
MR U. T. PLACE
DR R. H. PORTER
DR E. C. POULTON
DR A. PRESLY
DR P. RABBITT
DR J. T. REASON
MR G. ROCHFORD
MR D. SALTER
DR B. SEMEONOFF
DR L. H. SHAFFER
DR T. SHALLICE
DR M. H. SHELDON
MR J. W. SHEPHERD
DR R. SHUTER
DR A. J. SIMPSON
DR C. P. SPENCER
DR K. T. STRONGMAN
PROF. J. R. SYMONS
MR R. THOMSON
DR A. C. C. TREACHER
DR A. TREISMAN
PROF. M. TREISMAN
DR G. S. TUNE
PROF A. M. UTTLEY
MR J. VALENTINE
PROF. P. H. VENABLES
PROF. M. D. VERNON
MISS M. A. VINCE
PROF. D. VOWLES
DR P. C. WASON
DR J. WILCOX
DR J. WILDING
MISS M. M. WOODHEAD
MR D. S. WRIGHT
DR P. WRIGHT

CHANGE-OVER TO SI UNITS

From January 1972 the journals of the British Psychological Society will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Further information about SI units is contained in the revised *Suggestions to Authors* pamphlet issued by the Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

Authors who refer to physical measures in their paper should now normally use SI units; common units of time (e.g. hour, year) will, of course, persist. Conversion tables will be published during the change-over period, but in the *British Journal of Psychology* these will be restricted to units of length, as others rarely occur. When they do, conversion factors will be stated in footnotes.





Sorry about 'Softly Softly'
— it's one of the few programmes you cannot get with our modules
aim bioSCIENCES
Modular programming equipment for behavioural research
AIM bioSciences Ltd
The River Mill, St Ives, Huntingdon
(04806) 2901 telex 32223

International Journal of Psychology
Journal International de Psychologie

Quarterly

Published by the International Union of Psychological Science

The *International Journal of Psychology* is devoted to cross-cultural comparisons of psychological phenomena in general psychology, child psychology and social psychology. It also deals with international confrontation of theories and methods on which psychologists from various countries differ because of their native culture and training. Oriented towards basic research and theory, it publishes scientific papers in English or in French. In each issue, an International Platform gives news and information on internationally significant problems.

VOLUME 6 1971 No. 2

CONTENTS SOMMAIRE

THE BRITISH JOURNAL OF
PSYCHOLOGY

Edited by W. SLUCKIN

with J. ANNETT, W. S. ANTHONY, J. BROWN, R. CONRAD, N. F. DIXON, E. D. FRASER, J. A. GRAY, I. MARTIN, R. L. REID, H. R. SCHAFFER *and* P. B. WARR

Assistant Editor: DEREK JEHU
Book Reviews Editor: D. S. WRIGHT
Press Editor: ALBERT DICKSON

VOLUME 62, 1971

CAMBRIDGE
AT THE UNIVERSITY PRESS
1971

PUBLISHED BY
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W. 1
American Branch: 32 East 57th Street, New York, N.Y. 10022

Printed in Great Britain at the University Printing House, Cambridge

CONTENTS OF VOLUME 62

PART 1. FEBRUARY 1971

PAGE

Part 2. May 1971

Part 3. August 1971

PART 4. NOVEMBER 1971

INDEX OF AUTHORS

INDEX OF PUBLICATIONS REVIEWED

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology

The British Journal of Medical Psychology

The British Journal of Social and Clinical Psychology

The British Journal of Mathematical and Statistical Psychology

The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology, *The British Journal of Medical Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts, the subscription price, payable in advance, being £7.50 per volume post free ($25.000 in the U.S.A.) for *The British Journal of Psychology*, and £6.50 each ($22.00) for *The British Journal of Medical Psychology* and *The British Journal of Social and Clinical Psychology*. Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £6 per volume post free ($18 in the U.S.A.). Subscriptions should be sent to the Society's Agents, John Wright & Sons Ltd., 42–44 Triangle West, Bristol BS8 1EX; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year in February, June and November. The subscription is £4 a year post free, payable in advance, or £1.50 each number. The subscription price in the U.S.A. is $8.40. Subscriptions should be sent to the Subscription Dept Manager, Associated Book Publishers Ltd, Distribution Centre, North Way, Andover, Hants.

The series of *Monograph Supplements* issued in connexion with *The British Journal of Psychology* is now discontinued. Orders may be sent to any bookseller, or Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB from whom details of these publications may be obtained.

MANUSCRIPTS

or *The British Journal of Psychology* should be sent to

Professor W. Sluckin, Department of Psychology, The University, Leicester, LE1 7RH;

for *The British Journal of Medical Psychology* should be sent to

Dr J. J. Sandler, Room 101, Department of Psychiatry, Institute of Psychiatry, De Crespigny Park, London, S.E. 5;

for *The British Journal of Social and Clinical Psychology* should be sent to

Dr Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh EH8 9TJ; *or* to
Dr J. G. Ingham (Clinical Psychology Editor), M.R.C. External Staff, Llandough Hospital, Penarth, Glamorgan, CF6 1XW;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Dr A. R. Jonckheere, Department of Psychology, University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Professor J. D. Nisbet, Department of Education, University of Aberdeen, Old Aberdeen.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £8.50. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1.50 annually), *The British Journal of Medical Psychology* (£1.50 annually), *The British Journal of Social and Clinical Psychology* (£1.50 annually), *The British Journal of Mathematical and Statistical Psychology* (£1.50 annually), and *The British Journal of Educational Psychology* (£1.50 annually).
Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18–19 Albemarle Street, London, W1X4DN.

The British Journal of Psychology *Volume 62, Part 4, November 1971*

CONTENTS

Printed in Great Britain at the University Printing House, Cambridge